SPATIAL SCALING FOR THE NUMERICAL APPROXIMATION OF PROBLEMS ON UNBOUNDED DOMAINS

A Dissertation

by

DIMITAR VASILEV TRENEV

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2009

Major Subject: Mathematics

SPATIAL SCALING FOR THE NUMERICAL APPROXIMATION OF PROBLEMS ON UNBOUNDED DOMAINS

A Dissertation

by

DIMITAR VASILEV TRENEV

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	James Bramble
Committee Members,	Joseph Pasciak
	Raytcho Lazarov
	Vivek Sarin
Head of Department,	Al Boggess

December 2009

Major Subject: Mathematics

ABSTRACT

Spatial Scaling for the Numerical Approximation of
Problems on Unbounded Domains. (December 2009)
Dimitar Vasilev Trenev, B.S., Sofia University
Chair of Advisory Committee: Dr. James Bramble

In this dissertation we describe a coordinate scaling technique for the numerical approximation of solutions to certain problems posed on unbounded domains in two and three dimensions. This technique amounts to introducing variable coefficients into the problem, which results in defining a solution coinciding with the solution to the original problem inside a bounded domain of interest and rapidly decaying outside of it. The decay of the solution to the modified problem allows us to truncate the problem to a bounded domain and subsequently solve the finite element approximation problem on a finite domain.

The particular problems that we consider are exterior problems for the Laplace equation and the time-harmonic acoustic and elastic wave scattering problems.

We introduce a real scaling change of variables for the Laplace equation and experimentally compare its performance to the performance of the existing alternative approaches for the numerical approximation of this problem.

Proceeding from the real scaling transformation, we introduce a version of the perfectly matched layer (PML) absorbing boundary as a complex coordinate shift and apply it to the exterior Helmholtz (acoustic scattering) equation. We outline the analysis of the continuous PML problem, discuss the implementation of a numerical method for its approximation and present computational results illustrating its efficiency.

We then discuss in detail the analysis of the elastic wave PML problem and

its numerical discretization. We show that the continuous problem is well-posed for sufficiently large truncation domain, and the discrete problem is well-posed on the truncated domain for a sufficiently small PML damping parameter. We discuss ways of avoiding the latter restriction.

Finally, we consider a new non-spherical scaling for the Laplace and Helmholtz equation. We present computational results with such scalings and conduct numerical experiments coupling real scaling with PML as means to increase the efficiency of the PML techniques, even if the damping parameters are small.

To my parents, Васил and Анелия

ACKNOWLEDGMENTS

My sincerest gratitude goes first to my advisor, Prof. James Bramble. This dissertation would have never been possible without his inspiration, invaluable guidance and continuous support. Thank you, Dr. Bramble, for being there for me.

I would like to thank Prof. Raytcho Lazarov and Prof. Joseph Pasciak for their classes, which provided the basis for my research, and for the meetings and discussions, which helped me build my perception of numerical analysis. They both inspired and encouraged me throughout the years. I am eternally grateful to Prof. Lazarov for introducing me to Texas A&M and encouraging me to pursue a graduate degree here.

I would also like to specially thank Prof. Vivek Sarin for serving on my committee and for teaching an excellent course on Parallel and Distributed Numerical Algorithms.

I can not say enough to thank Prof. Bojan Popov and Prof. Guido Kanschat for providing me with financial support in the final year of my studies. I would especially like to thank Prof. Popov for his friendship and encouragement. He and his wife Rositza were my family away from home.

Finally, I would like to thank the faculty and staff at the TAMU Department of Mathematics for providing a great learning environment, and thank all my friends and officemates for making my time here even more enjoyable.

Thank you.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

In scattering problems, although one is generally interested in the solution close to the scatterer, if one wishes to straightforwardly apply the standard approximation methods, one would be forced to truncate the problem far enough from the scatterer to avoid large artificial reflections. This would lead to a huge linear system, which, for a real world three dimensional problem, even the more advanced supercomputers will be incapable of solving in a reasonable time. Thus, an absorbing boundary of some kind is usually used, which modifies the problem in such a way, that the solution in the region of interest stays (nearly) the same, while the distance to the artificial boundary (and hence the size of the computational domain) is greatly reduced.

In this dissertation, we propose a coordinate scaling technique to efficiently truncate particular problems posed on infinite domains to relatively small discrete problems. Contrary to the standard absorbing boundary conditions approach, this technique does not require complicated modifications of the problem since it amounts to introducing specific variable coefficients in the equation and using simple (e.g. zero Dirichlet) boundary conditions. As such, the resulting problem can be easily handled (in principle) by an existing finite element code (provided, in the case of the complex coordinate scaling, that the code allows the use of complex valued functions as coefficients).

The specific problems that we will look at are exterior problems for the Laplace, Helmholtz and elastic wave equations, which we will now state.

Let Ω be a closed bounded subset of $\mathbb{R}^d$ $(d = 2, 3)$ with a simply connected

---

This dissertation follows the style of the SIAM Journal on Numerical Analysis.

complement $\Omega^c = \mathbb{R}^d \setminus \Omega$. We shall consider the following model equations involving unknown functions defined on Ω^c.

- The exterior Laplace equation

$$\begin{aligned} -\Delta u &= 0 \quad \text{in} \quad \Omega^c, \\ u &= g_1 \quad \text{on} \quad \partial\Omega, \\ u &\to 0 \quad \text{as} \quad r \to \infty. \end{aligned} \tag{1.1}$$

- The exterior Helmholtz equation (the acoustic scattering problem)

$$\begin{aligned} -\Delta u - k^2 u &= 0 \quad \text{in} \quad \Omega^c, \\ u &= g_2 \quad \text{on} \quad \partial\Omega, \\ r^{\frac{d-1}{2}} \left(\frac{\partial u}{\partial r} - iku \right) &\to 0 \quad \text{as} \quad r \to \infty. \end{aligned} \tag{1.2}$$

  Here k is a positive real number, i is the imaginary unit and $r = \|\boldsymbol{x}\|_2$.

- The (frequency domain) elastic wave scattering problem

$$\begin{aligned} \boldsymbol{\Delta u} + \gamma \boldsymbol{\nabla} \nabla \cdot \boldsymbol{u} + k^2 \boldsymbol{u} &= \boldsymbol{0} \quad \text{in} \quad \Omega^c, \\ \boldsymbol{u} &= \boldsymbol{g} \quad \text{on} \quad \Gamma. \end{aligned} \tag{1.3}$$

  with the Kupradze-Sommerfeld radiation condition at infinity (see Section 4.1).

In Section 1.1, we illustrate the scaling technique on simple one dimensional examples and introduce the perfectly matched layer (PML) technique viewed as a complex coordinate shift.

Chapter II introduces the real space transformation, generalizing the example from Section 1.1 to two and three dimensions. We illustrate its application to the model Laplace problem (1.1) in two dimensions and present numerical experiments comparing it with the geometric approaches of mesh-grading and adaptive mesh refinement.

The PML is then introduced in Chapter III as a natural extension of that transformation to complex coordinates. There we rigorously define the complex coordinate shift, and discuss its application to the model Helmholtz problem in two dimensions (1.2). We then outline the analysis of the resulting problem and present some numerical results to illustrate the performance of the PML. Additionally, in Section 3.4, we discuss the finite element implementation and the specific issues arising in the discretization of the PML equation.

Chapter IV then presents in detail the analysis of the elastic wave PML problem. It is relatively self-contained, although some results from Chapter III will be required. The analysis of the discrete approximation of the elastic wave PML problem and some computational results are presented in Chapter V. The results from these two chapters will be published in [8].

In all chapters mentioned above, we use the distance from a point to the origin as a base for the scaling (i.e. spherical scaling). In Chapter VI, we briefly discuss a non-spherical transformation for both the real scaling and PML. Section 6.3 of this chapter also discusses the coupling of the real scaling and PML. Such an approach proves useful in speeding up the decay of of the solution when the damping parameter in the PML is small. As we shall see in the analysis of the numerical approximation for the elastic wave PML problem, a restriction (although seemingly artificial) on the size of the PML damping parameter is sometimes required to show well-posedness of the finite element PML problems. In Section 6.3 we present the results of different numerical experiments to draw some conclusions for the applicability of the aforementioned coupling.

Finally, Chapter VII serves as a summary of the dissertation and discusses possible future work.

1.1. One-dimensional examples

In this section, we illustrate the concepts of real coordinate scaling (change of variables) and complex coordinate scaling (PML) on simple one-dimensional problems coming from looking for radially symmetric solutions to the Laplace and Helmholtz problems in three dimensions (equations (1.1) and (1.2)). It serves as an introduction to some specific notation and terminology. The scaling technique will be introduced in more generality and detail in Chapter II. The PML complex coordinate shift is rigorously developed in Chapter III.

1.1.1. Real change of variables

Consider the second order ordinary differential boundary value problem for the unknown function $u(r) \in C^2([1,\infty))$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}u\right) = 0, u(1) = 1 \text{ and } \lim_{r\to\infty} u(r) = 0. \tag{1.4}$$

The differential operator here represents the radial part of the Laplacian (in spherical coordinates) and the above equation is equivalent to solving for a radially-symmetric harmonic function in three dimensions. Two linearly independent solutions of the differential equation (1.4) are $u_1(r) = \frac{1}{r}$ and $u_2(r) = 1$, and the decay condition at infinity identifies $u(r) = c_1 u_1(r)$ as our solution, with the value of the constant $c_1 = 1$ determined by the boundary condition at $r = 1$.

Suppose now that we are only interested in the solution close to the finite boundary of the domain of the equation, say in the interval $(1, n_0)$, which we shall call the "domain of interest". We would like to use the finite element method to approximate the solution of (1.4) there, and to do that we will need to truncate the equation to a finite computational domain. Since we are not interested in the solution for values

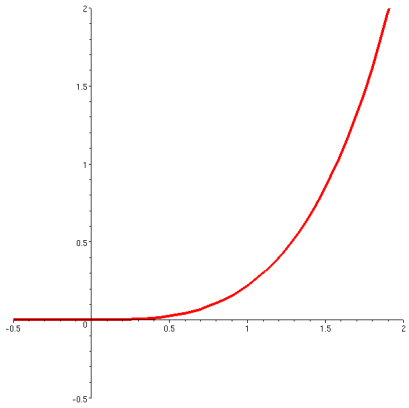

Fig. 1.1. Graph of the scaling function $\widetilde{\sigma}_{exp}(r)$; $\sigma_0 = 1$, $k = 2$.

of r greater than n_0, we can scale the variable in this part of the interval, effectively compressing the latter to a smaller computational domain.

To that end, we introduce the change of variables $r \to \widetilde{r} = \widetilde{d}(r)r$. We shall use a scaling factor of the form $\widetilde{d}(n) = 1 + \widetilde{\sigma}(n)$ (the reason for introducing both $\widetilde{d}$ and $\widetilde{\sigma}$ in the notation will become clear in the next section), where the function $\widetilde{\sigma}(n)$ is continuous, non-decreasing and equal to zero in the domain of interest (i.e. for $n \leq n_0$). This makes the change $r \to \widetilde{d}(r)r$ one-to-one and onto $(1, \infty)$, preserving the domain of interest (where we have $\widetilde{r} = r$).

In particular, for the real scaling, we shall use an exponentially increasing scaling function $\widetilde{\sigma}$ obtained by shifting (of n_0 units to the right) from function of the following form

$$\widetilde{\sigma}_{exp}(r) = e^{\sigma_0 r} - \sum_{j=0}^{k} \frac{(\sigma_0 r)^j}{j!}. \tag{1.5}$$

The parameters σ_0 and k in (1.5) allow us to control the growth and smoothness of the function respectively. A plot of such a function ($\sigma_0 = 1$, $k = 2$) is given in Figure 1.1

Equation (1.4) implies that the function $\widetilde{u}(r) = u(\widetilde{r})$ satisfies (in the interval

$(1,\infty)$) the equation

$$\frac{1}{\widetilde{r}^2}\frac{\partial}{\partial\widetilde{r}}\left(\widetilde{r}^2\frac{\partial}{\partial\widetilde{r}}\widetilde{u}\right)=0,$$

or (with $d(r)=\partial\widetilde{r}/\partial r$)

$$\frac{1}{\widetilde{d}^2r^2}\frac{1}{d}\frac{\partial}{\partial r}\left(\widetilde{d}^2r^2\frac{1}{d}\frac{\partial}{\partial r}\widetilde{u}\right)=0.$$

Imposing the boundary conditions, we can rewrite the equation for $\widetilde{u}$ as

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(\frac{\widetilde{d}^2}{d}r^2\frac{\partial}{\partial r}\widetilde{u}\right)=0,\widetilde{u}(1)=1 \text{ and } \lim_{r\to\infty}\widetilde{u}(r)=0. \tag{1.6}$$

Note that the exponential growth of $\widetilde{\sigma}(r)=\widetilde{\sigma}_{exp}(r-n_o)$ results in an exponential decay of the scaled solution $\widetilde{u}$ (and all its derivatives). This allows us to truncate equation (1.6) to a relatively small computational domain $\Omega_\infty=(1,R_t)$ imposing the boundary condition $\widetilde{u}_t(R_t)=0$. This of course, simulates truncating the unscaled equation (1.4) to the (significantly) larger domain $(1,\widetilde{R})$, where $\widetilde{R}=\widetilde{d}(R_t)R_t$, which means that the truncation error will be small. The steady decay in the derivatives of $\widetilde{u}$ ensures that this gain is not penalized by an increase in the discretization error and the coordinate scaling change is a viable approach for this problem.

To illustrate the performance of the scaling, assume our domain of interest is the interval $(1,3)$ and we want to achieve accuracy in our approximation of order 10^{-3}. In particular, we shall take $\widetilde{R}=2000$ (which guarantees that the truncation error is sufficiently small). Solving the unscaled equation on a uniform discretization of the finite computation domain $\widetilde{\Omega}_\infty=(1,2000)$, requires 262145 degrees of freedom to obtain the truncation error (i.e. the desired order of accuracy). On the other hand, taking $R_t=5$ and using a simple exponential scaling to simulate truncating at $\widetilde{R}=2000$ over the small computational domain $\Omega_\infty=(1,5)$ obtains the order of the truncation error in only 8193 degrees of freedom.

Figure 1.2 shows the plot of the exact solution $u(r)=\frac{1}{r}$ and the values of the

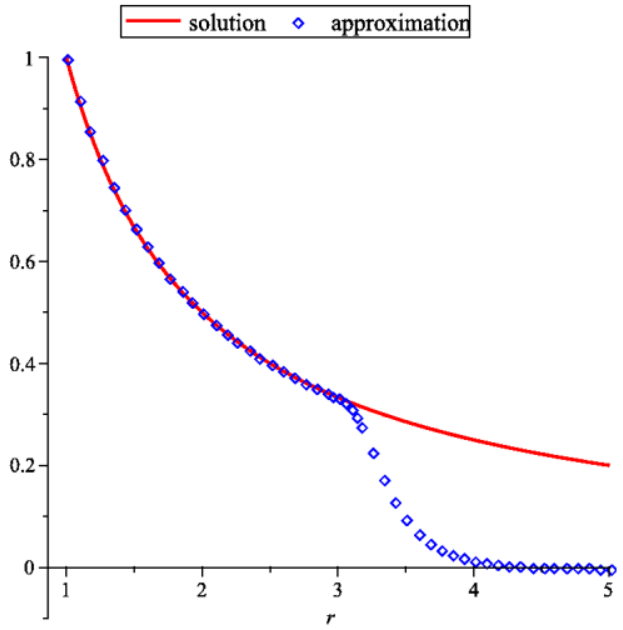

Fig. 1.2. Exact solution and its approximation.

computed scaled approximation at the endpoints of the (cubic) finite elements.

1.1.2. Complex coordinate shift (PML)

To illustrate the PML concept, consider the following boundary value problem posed on the interval $(1, \infty)$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}u\right) + k^2 u = 0, \tag{1.7}$$

defining a radially symmetric solution of the three dimensional Helmholtz equation. Here k is a positive real number.

Two linearly independent solutions of (1.7) are the functions $h_1(r) = \frac{e^{ikr}}{r}$ and $h_2(r) = \frac{e^{-ikr}}{r}$, and the general form of the solution is

$$u(r) = c_1 h_1(r) + c_2 h_2(r). \tag{1.8}$$

As both $h_1(r)$ and $h_2(r)$ are decaying at infinity we need a slightly more complicated

boundary condition than the one in (1.4) to make the equation well posed. Namely we impose the conditions

$$u(1) = e^{ik} \text{ and } \lim_{r\to\infty} r\left(u'(r) - iku(r)\right) = 0, \tag{1.9}$$

which identify $u(r) = h_1(r)$ as our solution. The boundary condition at infinity here is just the (three-dimensional) Sommerfeld radiation condition (see (1.2)).

While one can attempt to apply the exponential real scaling from the previous section in the numerical discretization of the above problem, this approach will no longer be as efficient. Indeed, the scaled solution $\widetilde{u}(r)$ becomes increasingly oscillatory in the compressed part of the region, and as a result (unlike the the previous case) its derivatives do not exhibit the same (exponential) decay. The increase in the higher seminorms of the scaled solution then needs to be compensated with a decrease in the mesh-size h (to reduce the overall discretization error), which significantly reduces the efficiency of the approach.

On the other hand, a scaling by a complex constant $r \to \widetilde{r} = (1 + i\sigma_0)r$, where σ_0 is a positive real number, will result in a negative real part of the exponent $ik\widetilde{r}$, once again introducing an exponential decay in the scaled solution ($\widetilde{u}(r) = h_1(\widetilde{r})$) and all of its derivatives. In addition, since with such a change of variables the function $h_2(\widetilde{r})$ becomes exponentially increasing, we no longer need the Robin boundary condition and may use a simple zero Dirichlet boundary condition to truncate to a finite computational domain.

Remark 1. *Note that while the definition $\widetilde{u}(r) = u(\widetilde{r})$ does not make sense for arbitrary functions $u : \mathbb{R} \mapsto \mathbb{C}$, it does make sense for functions of the form* (1.8)*, since both $h_1(r)$, $h_2(r)$ are well defined for complex values of the argument (excluding the branch cut of the negative real axis).*

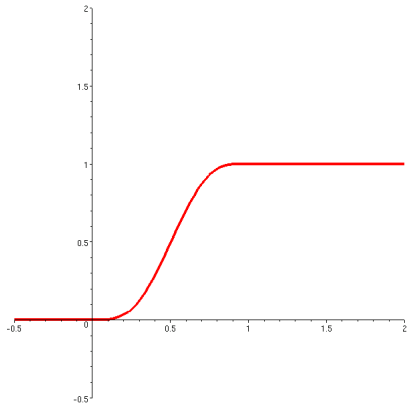

Fig. 1.3. Graph of the scaling function $\widetilde{\sigma}_{pml}(r)$; $a = 1$, $\sigma_0 = 1$, $k = 2$.

Since we want to preserve the solution in the domain of interest $(1, n_0)$ (and have $\widetilde{r} = r$ there), we gradually increase the imaginary part of the scaling factor by using $\widetilde{d}(r) = 1 + i\widetilde{\sigma}_{pml}(r - n_0)$. Here the function $\widetilde{\sigma}_{pml}(r)$ is given by the general form

$$\widetilde{\sigma}_{pml}(r) = \sigma_0 \frac{\int_0^{\max\{r,a\}} t^k (a-t)^k \, dt}{\int_0^a t^k (a-t)^k \, dt}. \tag{1.10}$$

A plot of such a function for particular values of the parameters a, σ_0 and k is given in Figure 1.3.

In view of Remark 1 we can formally change variables in equation (1.7) to derive the scaled (PML) equation analogously to equation (1.6). This results in the following equation for the function $\widetilde{u}(r)$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(\frac{\widetilde{d}^2}{d} r^2 \frac{\partial}{\partial r}\widetilde{u}\right) + k^2\widetilde{u} = 0, \widetilde{u}(1) = e^{ik} \text{ and } \lim_{r\to\infty} \widetilde{u}(r) = 0. \tag{1.11}$$

Here the function $d(r)$ is given (as in the previous section) as the derivative of $\widetilde{r}$ with respect to r.

Figure 1.4 shows the plot of the real part of the exact solution $u(r) = e^{ikr}$ and the real parts of the values of a computed PML approximation at the endpoints of the (cubic) finite elements. In this example the wavenumber was $k = 5$ and the "PML

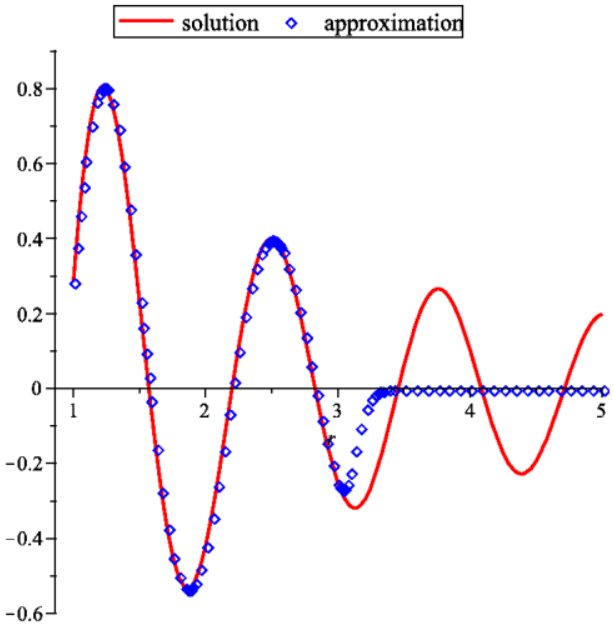

Fig. 1.4. Real part of the exact solution and its PML approximation.

damping parameter" σ_0 was taken $\sigma_0 = 0.5$. Note that the PML approximation is almost identical to the solution inside the domain of interest (here, as before, $n_0 = 3$) and quickly goes to zero outside of it.

CHAPTER II

REAL SPATIAL SCALING

In this chapter, we look at a real scaling transformation (change of variables), introducing some notation and setting the stage for the complex coordinate shift (PML) discussed in the next chapters. We will illustrate the application of this transformation to the Laplace equation and present the results of numerical experiments comparing it to existing solution techniques.

2.1. Notation

In general, throughout this dissertation, we will use bold symbols to denote vector valued functions, spaces and operators thus distinguishing them from the scalar valued ones. For example, in equations (1.1) and (1.2) the unknown u is a scalar function of the position $\boldsymbol{x}$, while $\boldsymbol{u}$ in equation (1.3) is a vector valued function. We shall denote the components of a vector or a vector valued function $\boldsymbol{w}$ by w_i, $i = 1 \dots d$.

In most of the dissertation, we shall have to deal with complex valued functions in various Sobolev spaces. For a domain D in $\mathbb{R}^d$, let $L^2(D)$ be the space of complex valued functions whose absolute values are square integrable on D and $\boldsymbol{L}^2(D) = (L^2(D))^d$ be the space of vector valued functions whose components are in $L^2(D)$. We shall use $(\cdot, \cdot)_D$ to denote the (vector or scalar Hermitian) $L^2(D)$-inner product and we shall omit the subscript, when there is no confusion about the domain of integration. The scalar and vector Sobolev spaces on D will be denoted $H^s(D)$ and $\boldsymbol{H}^s(D)$ respectively.

2.1.1. The scaling transformation

For the purpose of generality, instead of scaling based on the Euclidian distance of a point to the origin, we shall make use of a function $n : \mathbb{R}^d \mapsto \mathbb{R}^+$ describing the scaling parameter for a given point. For the most part one can think of $n(\boldsymbol{x}) = \|\boldsymbol{x}\|_2 = \left(\sum_j x_j^2\right)^{\frac{1}{2}}$, although we will consider different choices in Chapter VI.

We consider the space scaling transformation

$$\boldsymbol{T}_s\boldsymbol{x} = (1 + s\widetilde{\sigma}(n(\boldsymbol{x})))\boldsymbol{x} \equiv \widetilde{\boldsymbol{x}}, \tag{2.1}$$

where s is a complex number. In general the function $\tilde{\sigma} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ will be obtained by shifting one of the functions (1.5), (1.10) introduced in Section 1.1. The sole purpose of the general parameter s in (2.1) is the justification (by means of complex function theory) of some identities for the complex stretched operators (cf. Theorem 3 or equations (4.12)-(4.14)). In general, we will only be using $s = 1$, when talking about a real change of variables, and $s = i$ (the imaginary unit), when discussing PML.

Consistent with the notation in the examples of Section 1.1, we shall also use the functions $\widetilde{d}(n) = 1 + s\widetilde{\sigma}(n)$ and $d(n) = 1 + s\sigma(n)$, where $\sigma(n) = \widetilde{\sigma}(n) + n\widetilde{\sigma}'(n)$.

The derivative (Jacobian) of the change of variables (2.1) is

$$\boldsymbol{J}(\boldsymbol{x}) = \widetilde{d}(n(\boldsymbol{x}))\boldsymbol{I} + s\widetilde{\sigma}'(n(\boldsymbol{x}))n(\boldsymbol{x}) \begin{pmatrix} \frac{x_1}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_1} & \frac{x_1}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_2} & \frac{x_1}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_3} \\ \frac{x_2}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_1} & \frac{x_2}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_2} & \frac{x_2}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_3} \\ \frac{x_3}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_1} & \frac{x_3}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial y_2} & \frac{x_3}{n(\boldsymbol{x})}\frac{\partial n(\boldsymbol{x})}{\partial x_3} \end{pmatrix}, \tag{2.2}$$

or, denoting by $\boldsymbol{P}(\boldsymbol{x})$ the matrix above involving the partial derivatives of $n(\boldsymbol{x})$, we have

$$\boldsymbol{J}(\boldsymbol{x}) = \widetilde{d}(n(\boldsymbol{x}))\boldsymbol{I} + s\widetilde{\sigma}'(n(\boldsymbol{x}))n(\boldsymbol{x})\boldsymbol{P}(\boldsymbol{x}) = \widetilde{d}(\boldsymbol{x})(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{x})) + d(\boldsymbol{x})\boldsymbol{P}(\boldsymbol{x}). \tag{2.3}$$

To simplify the notation in the future, we will often exclude the variables in the function notation. Thus, we shall write equation (2.3), for example, as

$$\boldsymbol{J} = \widetilde{d}(n)\boldsymbol{I} + s\widetilde{\sigma}'(n)n\boldsymbol{P} = \widetilde{d}(\boldsymbol{I} - \boldsymbol{P}) + d\boldsymbol{P}. \tag{2.4}$$

2.1.2. Scaling of functions and operators

In this subsection (and only here), we shall look at two different copies of $\mathbb{R}^d$, containing D and $\widetilde{D}$ respectively, where in D we will denote the variable by $\boldsymbol{x}$ and in $\widetilde{D}$ with $\widetilde{\boldsymbol{x}}$. The underlying connection between the two variables will, of course, be the one given in (2.1).

For a given function $u \in C^\infty(\widetilde{D})$, one can define the corresponding function $\widetilde{u} : D \mapsto \mathbb{R}$ by

$$\widetilde{u}(\boldsymbol{x}) = u(\boldsymbol{T}_s\boldsymbol{x}) = u(\widetilde{\boldsymbol{x}}). \tag{2.5}$$

We shall also use "scaled" versions of differential operators. We define the operator $\widetilde{\boldsymbol{\nabla}}$ acting on $\widetilde{u}$ by

$$\widetilde{\boldsymbol{\nabla}}\,\widetilde{u} = \boldsymbol{J}^{-T}\boldsymbol{\nabla}\,\widetilde{u}. \tag{2.6}$$

Naturally, the idea behind this definition is to relate the gradients of u and $\widetilde{u}$. Note that the chain rule and definition (2.5) give

$$\boldsymbol{\nabla}\,\widetilde{u} = \boldsymbol{\nabla}\,u(\widetilde{\boldsymbol{x}}(\boldsymbol{x})) = \boldsymbol{J}^T\widetilde{\boldsymbol{\nabla}\,u}, \text{ i.e. } \widetilde{\boldsymbol{\nabla}}\,\widetilde{u} = \widetilde{\boldsymbol{\nabla}\,u}.$$

To relate the L^2- inner products on the two domains, note that we have, via a change of variables in the integrals, the identity

$$(u, v)_{\widetilde{D}} = \int_{\widetilde{D}} u(\widetilde{\boldsymbol{x}})v(\widetilde{\boldsymbol{x}})\,d\widetilde{\boldsymbol{x}} = \int_D \widetilde{u}(\boldsymbol{x})\widetilde{v}(\boldsymbol{x})\,(\det \boldsymbol{J}(\boldsymbol{x}))\,d\boldsymbol{x} = ((\det \boldsymbol{J})\widetilde{u}, \widetilde{v})_D. \tag{2.7}$$

We can now use the relations (2.7) and (2.6) to derive a formula for the stretched divergence operator (satisfying $\widetilde{\nabla} \cdot \widetilde{\boldsymbol{F}} = \widetilde{\nabla \cdot \boldsymbol{F}}$). For a given $\boldsymbol{F} \in \boldsymbol{C}^\infty(\widetilde{D})$ and an

arbitrary $\phi \in C_0^\infty(\widetilde{D})$, we have (by using integration by parts and the formulas above)

$$\begin{aligned}((\det \boldsymbol{J})\widetilde{\nabla} \cdot \widetilde{\boldsymbol{F}}, \widetilde{\phi})_D &= ((\det \boldsymbol{J})\widetilde{\nabla \cdot \boldsymbol{F}}, \widetilde{\phi})_D = (\nabla \cdot \boldsymbol{F}, \phi)_{\widetilde{D}} = -(\boldsymbol{F}, \boldsymbol{\nabla}\, \phi)_{\widetilde{D}} \\ &= -((\det \boldsymbol{J})\widetilde{\boldsymbol{F}}, \widetilde{\boldsymbol{\nabla}\, \phi})_D = -((\det \boldsymbol{J})\widetilde{\boldsymbol{F}}, \widetilde{\boldsymbol{\nabla}}\, \widetilde{\phi})_D \\ &= -((\det \boldsymbol{J})\widetilde{\boldsymbol{F}}, \boldsymbol{J}^{-T}\boldsymbol{\nabla}\, \widetilde{\phi})_D = (\nabla \cdot ((\det \boldsymbol{J})\boldsymbol{J}^{-1}\boldsymbol{F}), \widetilde{\phi})_D.\end{aligned}$$

Thus,

$$\widetilde{\nabla} \cdot \widetilde{\boldsymbol{F}} = \frac{1}{\det \boldsymbol{J}} \nabla \cdot ((\det \boldsymbol{J})\boldsymbol{J}^{-1}\widetilde{\boldsymbol{F}}). \tag{2.8}$$

The stretched Laplacian (satisfying $\widetilde{\Delta}\widetilde{u} = \widetilde{\Delta u}$) is then the natural composition of the above two operators, i.e.

$$\widetilde{\Delta} = \frac{1}{\det \boldsymbol{J}} \nabla \cdot ((\det \boldsymbol{J})\boldsymbol{J}^{-1}\boldsymbol{J}^{-T}\nabla). \tag{2.9}$$

Note that although derived for smooth functions, the formulas above define valid operators on Sobolev spaces. For example, (2.6) can be viewed both as an operator from $H^1(D)$ to $\boldsymbol{L}^2(D)$ and as an operator from $L^2(D)$ to $\boldsymbol{H}^{-1}(D)$.

2.1.3. Spherical scaling

For the particular choice $n(\boldsymbol{x}) = \|\boldsymbol{x}\|_2 = r$, we have the following formula for the matrix $\boldsymbol{P}$, involved in the Jacobian (2.4)

$$\boldsymbol{P} = \frac{1}{r^2} \begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 \\ x_1x_2 & x_2^2 & x_2x_3 \\ x_1x_3 & x_2x_3 & x_3^2 \end{pmatrix}. \tag{2.10}$$

In spherical coordinates, $\boldsymbol{J}$ is diagonal. Indeed, if $\boldsymbol{e}_r$, $\boldsymbol{e}_\phi$, and $\boldsymbol{e}_\theta$ are the spherical unit vectors and

$$\boldsymbol{u} = u_r\boldsymbol{e}_r + u_\phi\boldsymbol{e}_\phi + u_\theta\boldsymbol{e}_\theta,$$

then

$$\boldsymbol{J}\boldsymbol{u} = du_r\boldsymbol{e}_r + \widetilde{d}(u_\phi \boldsymbol{e}_\phi + u_\theta \boldsymbol{e}_\theta).$$

To make the notation conforming to the one in [6] and [8], we introduce the following matrices

$$\begin{aligned} \boldsymbol{A} = \frac{1}{\det \boldsymbol{J}}\boldsymbol{J} &= \frac{1}{\widetilde{d^2}}\boldsymbol{P} + \frac{1}{d\widetilde{d}}(\boldsymbol{I}-\boldsymbol{P}), \\ \boldsymbol{B} = \boldsymbol{J} &= d\boldsymbol{P}\boldsymbol{u} + \widetilde{d}(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{u}. \end{aligned} \tag{2.11}$$

With this notation we now have

$$\widetilde{\boldsymbol{\nabla}} w = \boldsymbol{B}^{-1}\boldsymbol{\nabla}\, w,\ \widetilde{\nabla}\cdot\boldsymbol{F} = \frac{1}{\widetilde{d^2}d}\nabla\cdot(\boldsymbol{A}^{-1}\boldsymbol{F}),\ \text{and}\ \widetilde{\Delta} w = \frac{1}{\widetilde{d^2}d}\nabla\cdot(\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}\boldsymbol{\nabla}\, w). \tag{2.12}$$

2.2. Real scaling for the Laplace problem - an example

In this section we shall illustrate the application of the above change of variables to the Laplace equation in two dimensions (1.1), which we restate as

$$\begin{aligned} -\Delta u &= f \quad \text{in} \quad \Omega^c, \\ u &= 0 \quad \text{on} \quad \partial\Omega, \\ u &\to 0 \quad \text{as} \quad \|\boldsymbol{x}\|_2 = r \to \infty. \end{aligned} \tag{2.13}$$

Here the function f is obtained from an appropriate extension of the boundary data of equation (1.1) and is taken to be compactly supported on the domain of interest.

Applying the change of variables, and denoting by $\boldsymbol{\mu}$ the matrix coefficient $(\det \boldsymbol{J})\boldsymbol{J}^{-1}\boldsymbol{J}^{-T}$, we can restate the above equation as

$$-\frac{1}{\det \boldsymbol{J}}\nabla\cdot\boldsymbol{\mu}\boldsymbol{\nabla}\,\widetilde{u} = \widetilde{f}, \tag{2.14}$$

with the same boundary conditions. The explicit form of the coefficient is

$$\boldsymbol{\mu} = \frac{\widetilde{d}}{d}\boldsymbol{P} + \frac{d}{\widetilde{d}}(\boldsymbol{I}-\boldsymbol{P}), \tag{2.15}$$

where $\boldsymbol{P}$ is the two-dimensional analogue of (2.10).

Note that outside a ball of radius r_0 containing the support of f, the solution of equation (2.13) has the series expansion

$$u(\boldsymbol{x}) = \sum_{n=1}^{\infty} \frac{1}{r^n} p_n(\hat{\boldsymbol{x}}), \tag{2.16}$$

here $r = |\boldsymbol{x}|$ and $\hat{\boldsymbol{x}} = \boldsymbol{x}/r$, and p_n are functions on the unit circle (of the form $a_n e^{in\theta} + b_n e^{-in\theta}$).

It is easy to see that, as in the one dimensional example in Section 1.1, the use of a scaling function $\widetilde{\sigma}(r) = \widetilde{\sigma}_{exp}(r - r_0)$ introduces exponential decay in the solution $\widetilde{u}(\boldsymbol{x})$ and all its derivatives, which suggests that the discretization error should not be significant factor in the numerical approximation.

We shall employ the finite element method to approximate the solution of (2.14) in the domain of interest ($\|\boldsymbol{x}\|_2 = r \leq r_0$). To that end, we first truncate the infinite domain equation to a finite computational domain $\Omega_\infty \supset B_{R_t} \setminus \Omega$, by imposing zero Dirichlet boundary condition on the outer boundary as well

$$\begin{aligned} -\tfrac{1}{\det \boldsymbol{J}} \nabla \cdot \boldsymbol{\mu} \boldsymbol{\nabla} \widetilde{u}_t &= \widetilde{f} \quad \text{in} \quad \Omega_\infty, \\ \widetilde{u}_t &= 0 \quad \text{on} \quad \partial\Omega_\infty. \end{aligned} \tag{2.17}$$

Integrating by parts, we obtain the following weak form of equation (2.17): find $\widetilde{u}_t \in H_0^1(\Omega_\infty)$ such that

$$(\boldsymbol{\mu} \boldsymbol{\nabla} \widetilde{u}_t, \boldsymbol{\nabla} v)_{\Omega_\infty} = ((\det \boldsymbol{J}) \widetilde{f}, v)_{\Omega_\infty} \text{ for all } v \in H_0^1(\Omega_\infty). \tag{2.18}$$

Note that generally the domain of interest contains the support of f, and so the function $\widetilde{\sigma}$ controlling the scaling is equal to zero, where f is not (and vice versa). In such cases, we simply have $(\det \boldsymbol{J}) \widetilde{f} \equiv f$.

On the continuous level the problem (2.18) is equivalent to solving the origi-

nal unscaled equation truncated to a domain $\widetilde{\Omega}_\infty = \{\widetilde{\boldsymbol{x}} \,|\, \boldsymbol{x} \in \Omega_\infty\}$. A change of variables on the discrete level shows that discretizing with piecewise polynomial functions $\{f \,|\, f|_K \in P_n(K)\}$ on a uniform mesh $\{K\}$ over the domain Ω_∞ is equivalent to discretizing the original variational problem with functions $\{f \,|\, f|_{\widetilde{K}} = p \circ \boldsymbol{T}_s^{-1}, p \in P_n(K)\}$ on the appropriately scaled mesh $\{\widetilde{K}\}$ over $\widetilde{\Omega}_\infty$.

As we shall see in the next section, there is little difference in the results obtained by solving the scaled problem (2.18) on a uniform mesh of Ω_∞ and those obtained by solving the original Laplace problem on the correspondingly graded mesh of $\widetilde{\Omega}_\infty$, even if we use the same (piecewise polynomial) finite element functions.

The (exponential) mesh grading for the numerical discretization of the exterior Laplace equation was proposed and analyzed in [16], where it was shown that the approximation error is quasi-optimal with respect to the number of degrees of freedom.

The real scaling approach described above is basically an algebraic way of implementing the (intrinsically geometric) mesh grading approach. It is somewhat easier to implement since it amounts to simply introducing a variable coefficient matrix $\boldsymbol{\mu}$ into the problem.

One last remark, we would like to make, is that the coefficient matrix $\boldsymbol{\mu}$ is "well-behaved" (from a computational point of view) in the sense that it does not exhibit steep growth (even if we use rapidly increasing scaling function $\widetilde{\sigma}$) or anisotropy in different directions. Indeed, we clearly have the bounds

$$\min_{\boldsymbol{x}} \frac{\widetilde{d}}{d}(\boldsymbol{\nabla}\, w, \boldsymbol{\nabla}\, w) \le (\boldsymbol{\mu}\boldsymbol{\nabla}\, w, \boldsymbol{\nabla}\, w) \le \max_{\boldsymbol{x}} \frac{d}{\widetilde{d}}(\boldsymbol{\nabla}\, w, \boldsymbol{\nabla}\, w), \tag{2.19}$$

and

$$\max_{\boldsymbol{x}} \frac{d(n(\boldsymbol{x}))}{\widetilde{d}(n(\boldsymbol{x}))} = \max_{\boldsymbol{x}} \left(1 + \frac{\widetilde{\sigma}'(n(\boldsymbol{x}))}{\widetilde{\sigma}(n(\boldsymbol{x}))} n(\boldsymbol{x})\right) \le 1 + C\sigma_0 \max_{\boldsymbol{x}} n(\boldsymbol{x}).$$

This shows that the eigenvalues of $\boldsymbol{\mu}$ grow (decrease) only linearly with σ_0 and R_t

(the size of Ω_∞) - quantities which are typically small compared to the simulated truncation radius $\widetilde{R}_t$ (we have $\sigma_0 R_t \approx \ln \widetilde{R}_t$).

2.2.1. Experiments and numerical results

In the numerical experiments conducted in this section, as well as most of the numerical results reported in the remainder of this dissertation, we work in the same test setting, which we will refer to as "standard" and which we describe here.

We consider a problem in two dimensions (of the type (1.1), (1.2) or (1.3)), where the scatterer Ω is a square centered at the origin with side length two units, i.e. $\Omega = [-1,1] \times [-1,1]$. The domain of interest is given by $\Omega_{int} = B(0,3) \setminus \Omega$, where $B(0,3)$ denotes the open disc of radius $r = 3$ centered at the origin. Finally, the computational domain Ω_∞ is $(-5,5) \times (-5,5) \setminus \Omega$.

The finite element discretization is done on a quadrilateral mesh over the domain Ω_∞ with bilinear (Q1), biquadratic (Q2) or bicubic (Q3) elements. In almost all cases the numerical experiments were realized by a C++ code written by the author in the framework of the deal.II finite element library (cf. [2]). Details of the implementation are discussed in Section 3.4 for the more complicated PML case, where we work with complex arithmetic and matrices, which are not necessarily symmetric and positive definite.

The particular experiments presented in this section are for equation (1.1) in the above setting. The boundary data on $\partial\Omega$ is taken so that the solution is $u(\boldsymbol{x}) = u(r,\theta) = \frac{\cos\theta}{r}$. The scaling function $\widetilde{\sigma}(r) = \widetilde{\sigma}_{exp}(r-3)$ is obtained from the general expression for the function $\widetilde{\sigma}_{exp}$ (1.5), by taking $k = 4$ (making $\widetilde{\sigma}$ four times continuously differentiable) and $\sigma_0 = 3.7861$ (giving, with $R_t = 5$, a truncation radius $\widetilde{R}_t \approx 8000$).

Figure 2.1 presents plots of the actual solution and a finite element approximation

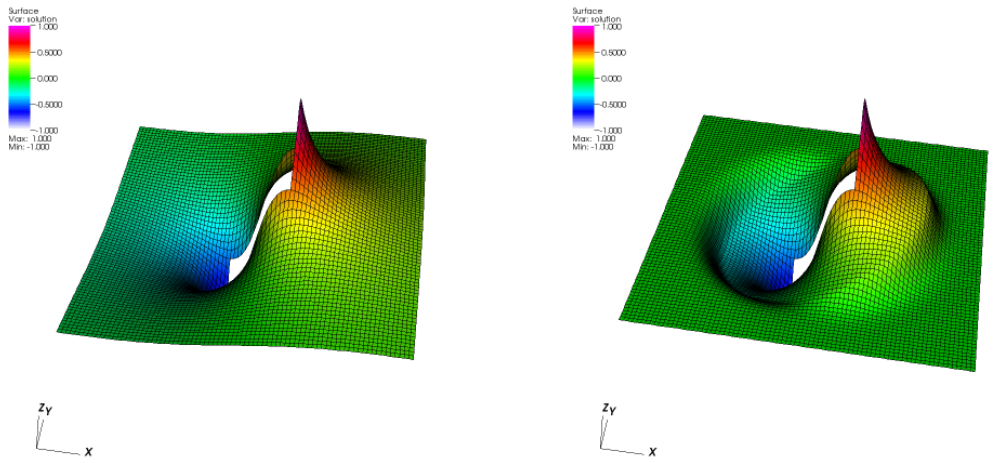

Fig. 2.1. The solution to the original problem and the scaled approximation.

of the scaled solution. Similar to the one dimensional case presented in Section 1.1, we see that the approximate solution is almost identical to the actual one in the domain of interest and goes to zero rapidly away from it.

We next compare the performance of the real scaling approach solving the unscaled problem (over the domain $(-8000, 8000)^2 \setminus \Omega$) using graded ([16]) or adaptively refined ([14]) meshes. Figure 2.2 shows an example of such meshes for the computational domain $(-20, 20) \times (-20, 20) \setminus \Omega$.

Not surprisingly in view of the discussion in the previous section, using a graded mesh, constructed by applying our scaling transformation to the vertices of a uniform mesh on the domain Ω_∞, produces results that are (after a few refinements) nearly identical to the ones obtained by solving the scaled problem on Ω_∞. In Table 2.1 we have presented the H^1-norm (in the domain of interest) of the error between the actual solution and its computed approximations.

We next turn to the adaptive refinement approach. The idea behind the adaptive mesh refinement is to use some computable quantity (*a-posteriori* error estimator) to

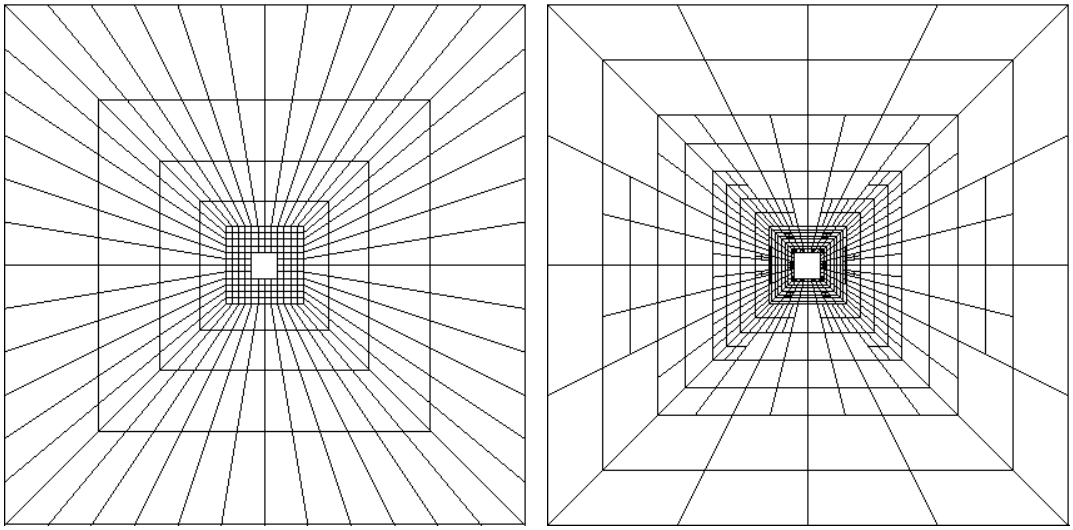

Fig. 2.2. Graded (left) and adaptively refined (right) meshes for the discretization of an exterior Laplace problem.

estimate the local error in the approximation and then refine the mesh only where needed (i.e. where the error is largest). This technique has proven to be an indispensable tool in the finite element computations in general, allowing for the location and satisfactory resolving of "rough areas" of the solution like singularities and discontinuities. However, its use in specific cases where the behavior of the solution is known *a-priori* is inefficient, since the computational cost of applying the error estimator can be avoided in such cases. We illustrate this, by comparing the real scaling approach to the results obtained on adaptively refined meshes. For such comparison, we start with the graded coarse mesh from the experiment above and perform several steps of adaptive refinement, using a standard error estimator by Kelly et. al ([18]). We report the step number (i.e. the number of consecutive refinements of the initial mesh) and the number of degrees of freedom required to obtain errors of comparable magnitude. The results of that experiment are presented in Table 2.2.

Table 2.1. H^1- norms of the error (graded meshes - left, real scaling - right)

cells	# dofs	graded mesh	rate	real scaling	rate
80	768	1.61626e-01	-	5.80627e-02	-
320	2976	2.70156e-02	2.58	9.48328e-03	2.61
1280	11712	1.85225e-03	3.87	8.85595e-04	3.42
5120	46464	1.19418e-04	3.96	9.65309e-05	3.20
20480	185088	1.24535e-05	3.26	1.19902e-05	3.01
81920	738816	1.51360e-06	3.04	1.50088e-06	3.00

Table 2.2. H^1- norms of the error and required d.o.f.s

real scaling			adaptive refinement		
step	# dofs	error	step	# dofs	error
0	768	5.80627e-02	2	3794	6.5846e-02
1	2976	9.48328e-03	5	37802	8.2865e-03
2	11712	8.85595e-04	7	170680	4.5246e-04
3	46464	9.65309e-05	8	355168	9.3923e-05
4	185088	1.19902e-05	10	1373244	1.0706e-05
5	738816	1.50088e-06	11	2595400	3.0924e-06

CHAPTER III

HELMHOLTZ PROBLEM - PML

In this chapter, we introduce the PML technique as a complex coordinate shift following the paper of Bramble and Pasciak [6]. We shall illustrate the PML applied to the model Helmholtz equation (1.2). The results of this chapter will be used in the analysis of the elastic wave PML equation in Chapter IV.

3.1. Introduction

A perfectly matched layer is an absorbing layer surrounding the computational domain, which causes no reflection at the interface between itself and the interior. The idea of a PML originates from a paper of Béreneger ([3]), where it was applied to a two dimensional version (transverse electric/magnetic modes) of the time-dependent Maxwell's equations. There it was viewed as a fictitious material with unphysical properties designed to absorb energy away from the domain of interest.

In [10] Chew and Weedon derived the model of [3] viewing the PML as a complex coordinate shift. This point of view made the formulation of the PML more convenient and easier to understand and analyze. It was further applied by Collino and Monk in [11] to derive a PML model, in rectangular and polar coordinates, for acoustic problems. A complete analysis of the spherical PML equations for the three dimensional electromagnetic and acoustic problems (both on the infinite and truncated domains) was given by Bramble and Pasciak in [6]. There they showed the well-posedness of both the infinite and finite domain PML problems as well as convergence of the solution of the truncated PML problem to the solution of the original infinite domain problem.

In this chapter, we shall follow [6] to introduce the PML model for the exterior

Helmholtz (acoustic scattering) equation and present a (slightly modified) proof of the *inf-sup* condition for the infinite domain PML problem. We shall need this result in the analysis of the elastic wave PML problem presented in the next chapter.

3.2. Preliminary results

In this section, we have quoted results that will be used in the analysis in the Helmholtz PML problem and (in case of Lemma 1) the elastic wave PML problem.

We start with a theorem that follows easily from a general result of Peetre [23] and Tartar [26] (see, e.g. Theorem 2.1 of [15]).

Theorem 1. *Let $A(\cdot,\cdot)$ be a bounded sesquilinear form on a (complex) Hilbert space V with norm $\|\cdot\|_V$. Let W be another Hilbert space with norm $\|\cdot\|_W$, and let $T : V \mapsto W$ be a compact operator. Suppose the following two conditions hold:*

(i) The only solution of

$$A(u,v) = 0 \text{ for all } v \in V$$

is $u = 0$.

(ii) There exist $C_1, C_2 > 0$ such that

$$\|u\|_V \le C_1 \sup_{v\in V} \frac{|A(u,v)|}{\|v\|_V} + C_2\|Tu\|_W \text{ for all } u \in V.$$

Then there exists $C_3 > 0$ such that for all $u \in V$,

$$\|u\|_V \le C_3 \sup_{v\in V} \frac{|A(u,v)|}{\|v\|_V} \tag{3.1}$$

We then state the generalized Lax-Milgram lemma.

Theorem 2. *Let H be a Hilbert space and let $\mathcal{B}$ be a bounded sesquilinear form on $H \times H$ satisfying, for $u \in H$,*

$$\|u\|_H \leq C \sup_{v \in H)} \frac{|\mathcal{B}(u,v)|}{\|v\|_H} \tag{3.2}$$

and

$$\|u\|_H \leq C \sup_{v \in H} \frac{|\mathcal{B}(v,u)|}{\|v\|_H}. \tag{3.3}$$

Then given $F \in H^*$ there exists a unique $u \in H$ satisfying

$$\mathcal{B}(u,v) = \langle F, v \rangle, \quad \textit{for all } v \in H.$$

Furthermore

$$\|u\|_H \leq C\|F\|_{H^*}.$$

We shall also state the following lemma, which is an immediate application of the preceeding Theorem.

Lemma 1. *Let D be a domain $\mathbb{R}^3$ with a Lipschitz boundary ∂D and let $\mathcal{B}$ be a bounded sesquilinear form on $H^1(D) \times H^1(D)$ satisfying, for $u \in H_0^1(D)$,*

$$\|u\|_{H^1(D)} \leq C \sup_{\phi \in H_0^1(D)} \frac{|\mathcal{B}(u,\phi)|}{\|\phi\|_{H^1(D)}} \tag{3.4}$$

and

$$\|u\|_{H^1(D)} \leq C \sup_{\phi \in H_0^1(D)} \frac{|\mathcal{B}(\phi,u)|}{\|\phi\|_{H^1(D)}}. \tag{3.5}$$

Then given $v \in H^{1/2}(\partial D)$ there exists a unique $u \in H^1(D)$ with $u = v$ on ∂D

satisfying

$$\mathcal{B}(u,\phi) = 0, \quad \textit{for all } \phi \in H_0^1(D).$$

Furthermore

$$\|u\|_{H^1(D)} \leq C\|v\|_{H^{1/2}(\partial D)}.$$

Remark 2. *In the case of the elastic wave equation, we shall use the above lemma for vector functions and spaces (*$\boldsymbol{u} \in \boldsymbol{H}^1(D)$, $\boldsymbol{v} \in \boldsymbol{H}^{1/2}(\partial D)$*).*

3.3. PML

The tilde operators of the previous chapter were defined on real valued functions and we extend them linearly to complex functions, e.g., if $w = u + iv$ with u and v real, then

$$\widetilde{\Delta}(u + iv) = \widetilde{\Delta}u + i\widetilde{\Delta}v.$$

Additionally, we extend these operators for values of the parameter s belonging to the complex right half-plane $Re\, z > -1/\sigma_M$, denoted by $\widetilde{\mathbb{C}}$ (here σ_M is the maximum value of the function $\sigma(r) = \widetilde{\sigma}(r) + r\widetilde{\sigma}'(r)$). This simply involves replacing s in $\widetilde{d}$ and d with z in the expressions (such as (2.6), (2.8) and (2.9)). Note that the change of integration variable formulas no longer make sense for z with nonzero imaginary part.

However, as we shall see in this section, one can still use those operators and apply the complex "scaling" to special functions such as solutions to the Helmholtz equation (1.2).

This section discusses equation (1.2), which we restate here

$$\begin{aligned} -\Delta u - k^2 u &= f \quad \text{in} \quad \Omega^c, \\ u &= 0 \quad \text{on} \quad \partial\Omega, \\ r^{\frac{d-1}{2}}\left(\tfrac{\partial u}{\partial r} - iku\right) &\to 0 \quad \text{as} \quad r \to \infty. \end{aligned} \tag{3.6}$$

We assume that the right hand side f, coming from applying the Helmholtz operator to an extension of the boundary data g_2 in (1.2), is compactly supported. Thus, outside a bounded ball $B_{r_{-1}}$ (containing the support of f), we have the following expansion for u:

$$u(\boldsymbol{x}) = \sum_{n=0}^{\infty} \sum_{|m|\le n} \alpha_{n,m} h_n^{(1)}(kr) Y_{n,m}(\hat{\boldsymbol{x}}), \tag{3.7}$$

Here $h_n^{(1)}$ is the Hankel function of the first kind of order n, $Y_{n,m}$ are spherical harmonics, $r = |\boldsymbol{x}|$ and $\hat{\boldsymbol{x}} = \boldsymbol{x}/r$. As is well known [13], the above series and all of its derivatives converge uniformly on compact subsets in the exterior of $\overline{B}_{r_{-1}}$.

In this section, we will work with a scaling function $\widetilde{\sigma}(r)$, obtained from $\widetilde{\sigma}_{pml}(r)$ through shifting. We will assume that the transition region (the region where $\widetilde{\sigma}(r)$ is not a constant) is contained in the interval (r_0, r_1), and that $r_0 > r_{-1}$ (i.e. $\widetilde{\sigma}(r) = 0$ for $r \le r_{-1}$ and we have not turned on the scaling inside $B_{r_{-1}}$).

Note that the Hankel functions in the expansion (3.7) have well defined values for complex arguments, provided that one stays away from the branch cut. For $z \in \widetilde{\mathbb{C}}$, the real part of $\tilde{r} = (1 + z\widetilde{\sigma}(r))r$ is positive, so we can take the branch cut defining $h_n^{(1)}$ to be the negative real axis. We can then define the stretched function $\widetilde{u}$ as

$$\widetilde{u}(\boldsymbol{x}) = \begin{cases} u(\boldsymbol{x}) & \text{if } |\boldsymbol{x}| \le r_0, \\ \displaystyle\sum_{n=0}^{\infty} \sum_{|m|\le n} \alpha_{n,m} h_n^{(1)}(k\widetilde{r}) Y_{n,m}(\hat{\boldsymbol{x}}) & \text{if } |\boldsymbol{x}| \ge r_0. \end{cases} \tag{3.8}$$

Then the function $\widetilde{u}$ defined above does indeed satisfy the expected "rescaled"

equation with the same right hand side, as stated in the following theorem.

Theorem 3. *Let z be in $\widetilde{\mathbb{C}}$. Then*

$$-\widetilde{\Delta}\widetilde{u} - k^2\widetilde{u} = f \quad \text{in } \Omega^c. \tag{3.9}$$

Proof. We consider the complex valued function

$$F(z) = (\det \boldsymbol{J}(\widetilde{\Delta}\widetilde{u} + k^2\widetilde{u}), \phi) + (f, \phi).$$

Here $\widetilde{u}$ depends on z as above and $\phi \in C_0^\infty(\Omega^c)$ is real valued and fixed. For real s with $s > -1/\sigma_M$, change of variables in the first integral gives

$$F(s) = (-f, \psi) + (f, \phi),$$

where $\psi(\boldsymbol{x}) = \phi(\boldsymbol{T}_s^{-1}\boldsymbol{x})$. As ϕ and ψ coincide on the open ball B_{r_0} and f is identically zero outside of it, we clearly have $F(s) = 0$. The series defining $\widetilde{u}$ converges uniformly on compact subsets with $|r| > r_{-1}$ from which it easily follows that F is analytic on $\widetilde{\mathbb{C}}$. This implies that $F(z) = 0$ for any $z \in \widetilde{\mathbb{C}}$. As ϕ was arbitrary, we conclude that (3.9) holds.

□

The PML shift is obtained by taking $z = i$ in the definition of the tilde operators of the previous chapter. This makes the definition of $\widetilde{u}$ and all of the tilde operators explicit. This choice shall be in effect in the remainder of the text.

We shall outline the analysis of the Helmholtz PML problem in two spatial dimensions. As we note in Remark 4, the considerations here extend directly to the three dimensional case.

We introduce the sesquilinear form

$$b(u, v) = (\boldsymbol{\mu}\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) - k^2(\widetilde{d}du, v). \tag{3.10}$$

Here the coefficient $\boldsymbol{\mu}$ is given as in (2.15) with the proper (complex) values for $\widetilde{d}$ and d.

Theorem 3 asserts (after multiplying the scaled equation by a test function and integrating by parts) that the function $w = \widetilde{u}$ satisfies the variational problem

$$b(w, \phi) = (\Psi, \phi) \text{ for all } \phi \in H_0^1(\Omega^c), \tag{3.11}$$

with data $\Psi = f$. We analyze the above problem and its adjoint

$$b(\phi, \hat{w}) = (\phi, \bar{\Psi}) \text{ for all } \phi \in H_0^1(\Omega^c). \tag{3.12}$$

Note that from the identity

$$b(u, v) = b(\overline{v}, \overline{u}) \tag{3.13}$$

it immediately follows that if w solves (3.11), then its conjugate $\hat{w} = \overline{w}$ solves the adjoint problem (3.12).

We will use Theorem 1 to show that the form $b(\cdot, \cdot)$ satisfies an *inf-sup* condition on $H_0^1(\Omega^c)$ (for both problems). Theorem 2 then asserts that the problems are well-posed. To begin we state a uniqueness result, whose proof is outlined in [11].

Proposition 1. *If $v \in H_0^1(\Omega^c)$ and v satisfies*

$$b(v, \phi) = 0 \text{ for all } \phi \in H_0^1(\Omega^c),$$

then $v = 0$.

We now prove a perturbed *inf-sup* condition, where the perturbation term is over the bounded domain $B_R \cap \Omega^c$. Here R is taken large enough ($R > r_1$) so that the ball B_R contains the PML transition region.

Proposition 2. *There exists $c > 0$ such that for any $u \in H_0^1(\Omega^c)$*

$$c\|u\|_{H_0^1(\Omega^c)} \leq \sup_{v \in H_0^1(\Omega^c)} \frac{|b(u,v)|}{\|v\|_{H_0^1(\Omega^c)}} + \|u\|_{L^2(B_R \cap \Omega^c)}. \tag{3.14}$$

Proof. In polar coordinates we have

$$\boldsymbol{\mu} = \begin{pmatrix} \frac{\widetilde{d}}{d} & 0 \\ 0 & \frac{d}{\widetilde{d}} \end{pmatrix} \text{ and } \Re(\boldsymbol{\mu}) = \begin{pmatrix} \frac{1+\widetilde{\sigma}\sigma}{1+\sigma^2} & 0 \\ 0 & \frac{1+\widetilde{\sigma}\sigma}{1+\widetilde{\sigma}^2} \end{pmatrix}.$$

Clearly $\Re(\boldsymbol{\mu})$ is a uniformly symmetric and positive definite matrix, and so

$$\Re((\boldsymbol{\mu}\boldsymbol{\nabla} u, \boldsymbol{\nabla} u)) \geq c_1(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u). \tag{3.15}$$

For a fixed σ_0, let $\alpha > \frac{1-\sigma_0^2}{2\sigma_0}$ and observe that past the PML transition layer (where $\widetilde{d} = d = 1 + i\sigma_0$), we have

$$\Re(-k^2(1+i\alpha)\widetilde{d}d) = \Re(-k^2(1+i\alpha)(1+i\sigma_0)^2) = -k^2(1-\sigma_0^2-2\alpha\sigma_0) > 0. \tag{3.16}$$

We shall prove (3.14) by constructing a particular $v \in H_0^1(\Omega^c)$, satisfying

$$c\|u\|_{H_0^1(\Omega^c)} \leq \frac{\Re(b(u,v))}{\|v\|_{H_0^1(\Omega^c)}} + \|u\|_{L^2(B_R \cap \Omega^c)}. \tag{3.17}$$

Denote by Ω_1 be the domain $B_{r_1} \cap \Omega^c$, and by Ω_R - the domain $B_R \cap \Omega^c$. Note that the PML transition region lies in Ω_1, and outside of it the coefficients in the form $b(\cdot,\cdot)$ are constant. Consider a sufficiently smooth function χ (whose real part is the constant 1), equal to 1 on Ω_1 and $1 + i\alpha$ on Ω_R^c. Note that $\boldsymbol{\nabla}\chi$ is non-zero only in the domain $T = \Omega_R \setminus \overline{\Omega}_1$.

We shall show that $v = \bar{\chi}u$ satisfies (3.17). To that end, we split the integrals in the expression $b(u,v)$ in several pieces as follows:

$$b(u,v) = b(u,\bar{\chi}u) = b_1(u,u) + b_2(u,u) + b_3(u,u) + l(u,u),$$

where

$$b_1(u,u) = (\boldsymbol{\mu}\boldsymbol{\nabla} u, \boldsymbol{\nabla} u)_{\Omega_1} + (u,u)_{\Omega_1},$$

$$b_2(u,u) = (\chi\boldsymbol{\nabla} u, \boldsymbol{\nabla} u)_T + (u,u)_T,$$

$$b_3(u,u) = ((1+i\alpha)\boldsymbol{\nabla} u, \boldsymbol{\nabla} u)_{\Omega_R^c} - k^2((1+i\alpha)\widetilde{d}du, u)_{\Omega_R^c},$$

and

$$l(u,u) = (\boldsymbol{\nabla} u, u\boldsymbol{\nabla}\,\bar{\chi})_T - k^2(\chi\widetilde{d}du, u)_{\Omega_R} - (u,u)_{\Omega_R}.$$

Note that (3.15) gives

$$c\|u\|_{H_0^1(\Omega_1)} \leq \Re(b_1(u,u)). \tag{3.18}$$

Using $\Re(\chi) = 1$, we clearly have

$$c\|u\|_{H_0^1(T)} \leq \Re(b_2(u,u)). \tag{3.19}$$

Finally, from (3.16) we obtain

$$\Re(-k^2((1+i\alpha)\widetilde{d}d\ u, u)_{\Omega_R^c}) \geq c_2(u,u)_{\Omega_R^c},$$

which gives

$$c\|u\|_{H_0^1(\Omega_R^c)} \leq \Re(b_3(u,u)). \tag{3.20}$$

Combining (3.18), (3.19) and (3.20) we obtain

$$c\|u\|^2_{H_0^1(\Omega^c)} = c(\|u\|^2_{H_0^1(\Omega_1)} + \|u\|^2_{H^1(T)} + \|u\|^2_{H^1(\Omega_R^c)}) \leq \Re(b_1(u,u) + b_2(u,u) + b_3(u,u)), \tag{3.21}$$

which, along with the triangle inequality and the bound

$$|\Re(l(u,u))| \leq C\|u\|_{H^1(\Omega_R)}\|u\|_{L^2(\Omega_R)} \leq C\|u\|_{H_0^1(\Omega^c)}\|u\|_{L^2(\Omega_R)},$$

gives us

$$c\|u\|^2_{H^1_0(\Omega^c)} \leq |\Re(b(u,\bar{\chi}u))| + C\|u\|_{H^1_0(\Omega^c)}\|u\|_{L^2(\Omega_R)}.$$

Since we clearly have $|\Re(b(u,\bar{\chi}u))| = |\Re(b(u,v))| \leq |b(u,v)|$, we can rewrite the last inequality as

$$c\|u\|_{H^1_0(\Omega^c)} \leq \frac{|b(u,v)|}{\|u\|_{H^1_0(\Omega^c)}} + C\|u\|_{L^2(\Omega_R)}. \tag{3.22}$$

Observing that the function χ and its reciprocal (and their conjugates) are bounded in $W^{1,\infty}(\mathbb{R}^2)$, we have $\|u\|_{H^1_0(\Omega^c)} \approx \|v\|_{H^1_0(\Omega^c)}$, and (3.17) immediately follows from (3.22). □

Remark 3. *For the spherical PML scaling, the result of Proposition 2 was proven in [6] as part of Theorem 3.1. The modified version of the proof, presented above, proves the result for any $\boldsymbol{\mu}$ with a symmetric and positive definite real part. In particular, we will refer to it in the analysis outline for a non-spherical PML in Chapter VI.*

We can now state the main result

Theorem 4. *Let Ψ be in $L^2(\Omega^c)$. Problems* (3.11) *and* (3.12) *have unique solutions $w, \hat{w} \in H^1_0(\Omega^c)$ satisfying*

$$\|w\|_{H^1(\Omega^c)} \leq C\|\Psi\|_{L^2(\Omega^c)} \text{ and } \|\hat{w}\|_{H^1(\Omega^c)} \leq C\|\Psi\|_{L^2(\Omega^c)}. \tag{3.23}$$

In addition, w and $\hat{w}$ are in $H^2(D)$, for any D, with $\overline{D} \subset \Omega^c$, and

$$\|w\|_{H^2(D)} \leq C\|\Psi\|_{L^2(\Omega^c)} \text{ and } \|\hat{w}\|_{H^2(D)} \leq C\|\Psi\|_{L^2(\Omega^c)}. \tag{3.24}$$

Proof. Let $T : H^1(\Omega^c) \mapsto L^2(\Omega_R)$ denote the $L^2(\Omega_R)$-embedding of restriction of elements of $H^1(\Omega^c)$. Note that T is a compact operator, since the domain Ω_R is bounded. Using Propositions 1 and 2, we can apply Theorem 1 with the above T to

conclude

$$c\|u\|_{H_0^1(\Omega^c)} \leq \sup_{v\in H_0^1(\Omega^c)} \frac{|b(u,v)|}{\|v\|_{H_0^1(\Omega^c)}}. \tag{3.25}$$

The corresponding *inf-sup* condition for the adjoint problem follows immediately from (3.13), and we can apply Theorem 2 to complete the proof of the first part of the theorem.

The proof of the second part, the interior regularity of the solutions, is standard. For ease of presentation we outline it here for the case when the domain of the problem Ω^c is the whole space $\mathbb{R}^2$. The general case is proven in the same manner with the help of a cut-off function, which is identically one in D and is supported away from the boundary of Γ of Ω^c.

Since $\widetilde{\sigma} \in C^2(\mathbb{R}^+)$ and $\sigma \in C^1(\mathbb{R}^+)$ we see that

$$|b(D_h v, \phi) + b(v, D_{-h}\phi)| \leq C\|v\|_{H^1(\mathbb{R}^2)}\|\phi\|_{H^1(\mathbb{R}^2)}, \tag{3.26}$$

where D_h is an arbitrary difference quotient of size h. Using (3.25) (with Ω^c replaced by $\mathbb{R}^2$) gives

$$\begin{aligned}
\|D_h w\|_{H^1(\mathbb{R}^2)} &\leq C\left(\sup_{\phi\in C_0^\infty(\mathbb{R}^2)} \frac{|b(w, D_{-h}\phi)|}{\|\phi\|_{H^1(\mathbb{R}^2)}} + \|u\|_{H^1(\mathbb{R}^2)} \right) \\
&= C\left(\sup_{\phi\in C_0^\infty(\mathbb{R}^2)} \frac{|(\Psi, D_{-h}\phi)|}{\|\phi\|_{H^1(\mathbb{R}^2)}} + \|u\|_{H^1(\mathbb{R}^2)} \right) \\
&\leq C\|\Psi\|_{L^2(\mathbb{R}^2)},
\end{aligned} \tag{3.27}$$

uniformly in h. For the last inequality above, we used (cf., Lemma 8.48 [24])

$$\|D_{-h}\phi\|_{L^2(\mathbb{R}^2)} \leq C\|\phi\|_{H^1(\mathbb{R}^2)}.$$

It follows that

$$\|w\|_{H^2(\mathbb{R}^2)} \leq C\|\Psi\|_{L^2(\mathbb{R}^2)} \tag{3.28}$$

(see, e.g., Lemma 8.49 of [24]).

The proof for $\hat{w}$ is the same.

□

Remark 4. *The results of this section are valid for the three dimensional case using the form*

$$b_{3D}(u,v) = (\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}\boldsymbol{\nabla}\, u, \boldsymbol{\nabla}\,(v/\overline{d})) - k^2(\widetilde{d}^2 du, (v/\overline{d})), \tag{3.29}$$

with the respective (complex) 3×3*-matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *given as in* (2.11).

The PML approximation is then defined, truncating the infinite-domain PML equation (3.9) to a finite domain $\Omega_\infty \subset \Omega^c$ by imposing zero Dirichlet boundary conditions on the (outer) boundary. That is to say, we are looking for a function $\widetilde{u}_t$ satisfying

$$\begin{aligned} -\widetilde{\Delta}\widetilde{u}_t - k^2\widetilde{u}_t &= f \quad \text{in} \quad \Omega_\infty, \\ \widetilde{u} &= 0 \quad \text{on} \quad \partial\Omega_\infty. \end{aligned} \tag{3.30}$$

The fact that the truncated PML problem is well-posed for a sufficiently large domain Ω_∞ and its solution converges exponentially (with the diameter of Ω_∞) to the solution $\widetilde{u}$ of the infinite domain PML problem was originally proved by Bramble and Pasciak in [6]. This result will follow immediately from the results proven in Chapter IV for the more general case of the elastic wave equation and we shall not discuss it further here. We shall, however, state the following two results, used in the analysis, which we will refer to when we consider the elastic wave problem.

The first is a classical interior estimate for the solution of an elliptic equation whose proof is elementary.

Proposition 3. *Suppose that* w *satisfies the equation*

$$\Delta w + \beta w = 0 \tag{3.31}$$

in a domain D *with a (possibly complex) constant* β*. If* D_1 *is a subdomain, whose*

closure is contained in D, then

$$\|w\|_{H^2(D_1)} \leq C\|w\|_{L^2(D)}. \tag{3.32}$$

The second is a decay estimate for the solution of the Helmholtz PML equation (cf. Proposition 3.2 of [6]).

Proposition 4. *Let $0 < a < b$ be fixed positive numbers and let $R_t > b$ be arbitrary. Denote by B_a, B_b and B_{R_t} the open balls centered at the origin with the respective radii.*

Assume that w is bounded at infinity and satisfies (3.31) *in $\Omega^c \setminus \overline{B}_a$ with $\beta = k^2(1 + i\sigma_0)^2$. Let $D \supset B_{R_t}$ be a domain in $\mathbb{R}^3$ and set $S_\alpha = \{\boldsymbol{x} \ : \ dist(\boldsymbol{x}, \partial D) < \alpha\}$, with $\alpha > 0$ fixed (independent of R_t) and small enough that $\overline{S}_\alpha \cap \overline{B}_b = \emptyset$.*

Then

$$\|w\|_{L^2(S_\alpha)} \leq Ce^{-\sigma_0 k R_t}\|w\|_{L^2(B_b \setminus \overline{B}_a)},$$

with a constant $C > 0$ independent of R_t.

3.4. Implementation and numerical results

In this section we shall discuss the issues arising in the implementation of the PML technique described above. As in the case of the real scaling, the PML technique amounts to introducing a variable coefficient in the equation. However, there are two issues that need to be addressed. First, we are now dealing with complex coefficients and the finite element code should be able to handle that. Second, by the nature of the PML coefficient, the resulting finite element matrices are no longer Hermitian and one should choose an appropriate solving technique.

3.4.1. Implementation

To begin with, we start with an explanation of the implementation in an abstract setting. Let X be a given Hilbert space over $\mathbb{C}$. We will denote by $(\cdot,\cdot)_X$ the inner product on X and assume it to be linear with respect to the first argument (and consequently conjugate-linear with respect to the second).

We shall consider a variational problem similar to (3.11) posed on the space X.

Let X^* denote the space of bounded conjugate-linear functionals on X. For $f \in X^*$ and $v \in X$ we introduce the "dual pairing" using the notation

$$\langle f, v\rangle = f(v) \text{ and } \langle v, f\rangle = \overline{\langle f, v\rangle}.$$

Note that any bounded sesquilinear (again, for consistency, linear with respect to the first argument) form $\mathcal{A} : X \times X \mapsto \mathbb{C}$ gives rise to a bounded linear map $\mathcal{L} : X \mapsto X^*$ defined by

$$\langle \mathcal{L}u, v\rangle = \mathcal{A}(u, v), \text{ for all } u, v \in X.$$

The adjoint (in the usual sense) of such map is then $\mathcal{L}^* : (X^*)^* \equiv X \mapsto X^*$ and is given by:

$$\langle \mathcal{L}^*u, v\rangle = \langle u, \mathcal{L}v\rangle = \overline{\langle \mathcal{L}v, u\rangle} = \overline{\mathcal{A}(v, u)}.$$

Clearly, if the form $\mathcal{A}(\cdot,\cdot)$ is Hermitian we have $\mathcal{L} \equiv \mathcal{L}^*$.

We will consider the variational problem: For a given $f \in X^*$, find $u \in X$ such that

$$\mathcal{A}(u, v) = \langle f, v\rangle \text{ for all } v \in X, \tag{3.33}$$

and more precisely its restriction to a finite-dimensional sub-space $X_h \subset X$.

For a given basis $\{\phi_i\}$ of X_h, the associated dual basis for X_h^* will (unless otherwise specified) be $\{\psi_i\}$ uniquely defined by the relation $\psi_i(\phi_j) = \delta_{ij}$.

Note that both X_h and X_h^* are isomorphic to $\mathbb{C}^n$ with $n = \dim X$. For $v \in X_h$ or $f \in X_h^*$, we will denote by $\vec{v} \in \mathbb{C}^n$ or $\vec{f} \in \mathbb{C}^n$ the vector of the coordinates of v or f in the corresponding basis.

With this notation $\langle f, v\rangle = (\vec{f}, \vec{v})_{\mathbb{C}^n} = \vec{v}^* \vec{f}$, where the * denotes the usual conjugate transpose.

It is easy to see how the operators corresponding to the form discussed above translate to matrices in $\mathbb{C}^{n\times n}$. For a given operator $\mathcal{L}_h : X_h \mapsto X_h^*$, we can consider the matrix L with $L_{i,j} = \langle \mathcal{L}\phi_j, \phi_i\rangle$. We then have the relation $\vec{\mathcal{L}_h u} = L\vec{u}$.

Note that the matrix L^* corresponding to $\mathcal{L}_h^* : (X_h^*)^* \equiv X_h \mapsto X_h^*$ is just the conjugate transpose of the matrix L (which justifies the use of the * notation). Indeed,

$$\langle \mathcal{L}_h^* u, v\rangle = \langle u, \mathcal{L}_h v\rangle = \overline{\langle \mathcal{L}_h v, u\rangle} = \overline{\vec{u}^* L \vec{v}} = \vec{v}^* L^* \vec{u}. \tag{3.34}$$

The variational problem (3.33), restricted to the subspace X_h, then gives rise to the matrix vector equation

$$L\vec{u} = \vec{f}, \tag{3.35}$$

since the approximate solution $u = \sum_i \vec{u}_i \phi_i$ will satisfy the equation in (3.33) for all $v \in X_h$.

The complex quantities involved can be broken in real and imaginary parts by using a standard bijective map $\mathcal{R} : \mathbb{C}^n \mapsto \mathbb{R}^{2n}$, given by

$$\mathcal{R}\vec{w} = \begin{pmatrix} \Re\vec{w} \\ \Im\vec{w} \end{pmatrix}.$$

We will also abuse notation and use $\mathcal{R}$ for mapping matrices. Given a matrix $L =$

$L_r + iL_i$, where L_r and L_i are real matrices (i.e. $L_r = \Re R$, $L_i = \Im L$), we define

$$\mathcal{R}L = \begin{pmatrix} L_r & -L_i \\ L_i & L_r \end{pmatrix}.$$

The reason for that definition is obvious, as with it we have the equality

$$\mathcal{R}(L\vec{w}) = \mathcal{R}(L)\mathcal{R}\vec{w},$$

for arbitrary $\vec{w} \in \mathbb{C}^n$.

This mapping also has the following convenient properties:

1. $\mathcal{R}(A^*) = \mathcal{R}(A)^t$ (and consequently the image of a Hermitian matrix is a symmetric one).

2. $\mathcal{R}(AB) = \mathcal{R}(A)\mathcal{R}(B)$.

3. The set of eigenvalues of $\mathcal{R}(A)$ consists of the eigenvalues of A and their conjugates.

We can now solve equation (3.35) using only real arithmetic, by solving the corresponding equation

$$\mathcal{R}(L)y = \mathcal{R}\vec{f}$$

for the unknown $y = \mathcal{R}\vec{u} \in \mathbb{R}^{2n}$.

Remark 5. *We compute the finite element approximation of the solution to the truncated PML equation* (3.30) *in the above abstract setting. We take X to be the space $H_0^1(\Omega_\infty)$, X_h - the finite element subspace of X, and the form $\mathcal{A}(\cdot,\cdot)$ to be the restriction the form $b(\cdot,\cdot)$, given in* (3.10)*, to the domain Ω_∞.*

It should be noted that the (real) matrices L_r and L_i can be computed in the usual

way from the bilinear forms

$$b_r(u,v) = (\Re(\boldsymbol{\mu})\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) - k^2(\Re(\widetilde{d}d)u, v)$$

and

$$b_i(u,v) = (\Im(\boldsymbol{\mu})\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) - k^2(\Im(\widetilde{d}d)u, v),$$

and so no complex arithmetic is needed anywhere in the process.

3.4.2. Negative-norm least squares approach

As we noted in the beginning of this section, the sesquilinear form (3.10) in the weak form of the Helmholtz PML equation is not Hermitian. As a result, the matrix resulting from the finite element discretization (i.e. the matrix $\mathcal{R}(L)$ resulting from taking $X = H_0^1(\Omega_\infty)$ and $\mathcal{A}(\cdot,\cdot) = b(\cdot,\cdot)$ in the above abstract setting and X_h - the finite element subspace of X) is not symmetric.

Thus, instead solving the problem (3.33), i.e.

$$\langle \mathcal{L}u, v\rangle = \langle f, v\rangle,$$

we shall solve the equivalent problem

$$(\mathcal{L}u, \mathcal{L}v)_{X^*} = (f, \mathcal{L}v)_{X^*}. \tag{3.36}$$

The inner-product on the space X^* can be related to the inner product on X by introducing the operator $\mathcal{T} : X^* \mapsto X$, defined by

$$(\mathcal{T}f, v)_X = \langle f, v\rangle, \text{ for given } f \in X^* \text{ and every } v \in X.$$

The natural norm on X^* is given by

$$\|f\|_{X^*} = \sup_{v\in X} \frac{|\langle f, v\rangle|}{\|v\|_X} = \sup_{v\in X} \frac{\|(\mathcal{T}f, v)_X\|}{\|v\|_X} = \|\mathcal{T}f\|_X.$$

The polarization formula and the above equality give the following formula for the X^* inner product

$$(f,g)_{X^*} = (\mathcal{T}f, \mathcal{T}g)_X. \tag{3.37}$$

Note that by the definition of $\mathcal{T}$

$$(f,g)_{X^*} = \langle f, \mathcal{T}g\rangle,$$

and it can be easily seen that $\mathcal{T}$ is self-adjoint

$$\langle f, \mathcal{T}g\rangle = (f,g)_{X^*} = \overline{(g,f)_{X^*}} = \overline{\langle g, \mathcal{T}f\rangle} = \langle \mathcal{T}f, g\rangle.$$

In the implementation we shall need the matrix T corresponding to the discrete version, $\mathcal{T}_h$, of the operator $\mathcal{T}$. Denote by D the matrix with elements $D_{ij} = (\phi_j, \phi_i)_{X_h}$ i.e. the matrix corresponding to the inner product on X_h. For any $u, v \in X_h$ one has

$$\vec{v}^* D\vec{u} = (u,v)_{X_h},$$

or, in particular, for a given $f \in X_h^*$ and for every $v \in X_h$

$$\vec{v}^* D\vec{\mathcal{T}_h f} = (\mathcal{T}_h f, v)_{X_h} = \langle f, v\rangle = \vec{v}^*\vec{f}.$$

The above equation implies $T\vec{f} = \vec{\mathcal{T}_h f} = D^{-1}\vec{f}$, or $T = D^{-1}$.

Equation (3.37) then gives the following formula for the inner product on X_h^*

$$(f,g)_{X_h^*} = (\mathcal{T}_h f, \mathcal{T}_h g)_{X_h} = (D^{-1}\vec{g})^* D(D^{-1}\vec{f}) = \vec{g}^* D^{-1^*}\vec{f} = \vec{g}^* D^{-1}\vec{f}.$$

In particular,

$$(\mathcal{L}_h u, \mathcal{L}_h v)_{X_h^*} = \vec{\mathcal{L}_h v} D^{-1} \vec{\mathcal{L}_h u} = \vec{v} L^* D^{-1} L\vec{u}, \tag{3.38}$$

and the discrete version of equation (3.36) reduces to the matrix vector equation

$$L^* D^{-1} L\vec{u} = L^* D^{-1}\vec{f}. \tag{3.39}$$

Remark 6. *Clearly since D and L are invertible matrices, equations* (3.35) *and* (3.39) *are equivalent. The reason we went through the derivation of* (3.39) *as a matrix-vector from of equation* (3.36) *is because we wanted to comment on the implementation of the negative-norm least-squares method, which we will also mention when discussing the elastic wave equation.*

Remark 7. *A discrete* inf-sup *condition, with a constant independent of the mesh size parameter, can be proven in a manner similar to the proof of* (3.25)*. Such a result guarantees that the number of iterations, when applying a preconditioned conjugate gradient (PCG) method to* (3.39) *with T as a preconditioner, will be bounded independently of the dimension of the discrete finite element space. Thus, the technique described above is a feasible approach for solving the equation resulting from the discretization of* (3.11)*.*

3.4.3. Computational results

We will now illustrate the performance of the PML method on a test problem. The exact solution u will be given by

$$u(\boldsymbol{x}) = u(r, \theta) = H_1^{(1)}(2r)\cos\theta.$$

Away from the origin u solves the homogeneous Helmholtz equation (with $k = 2$) and consequently satisfies equation (1.2) with the appropriate boundary data for any $\Omega \ni (0, 0)$. As in Section 2.2.1, we shall take $\Omega = [-1, 1]^2 \subset \mathbb{R}^2$ and be interested in the solution to our model equation only in the domain of interest $B(0, 3) \cap \Omega^c$. We define the PML approximation explicitly, by giving concrete values to the PML parameters. For the PML layer, we take $r_0 = 3$, $r_1 = 4$ and $\sigma_0 = 1$. For the finite element approximation, we truncate the computational domain to $\Omega_\infty = (-5, 5)^2 \setminus \overline{\Omega}$

and use bilinear finite elements over a uniform mesh (grid) of squares.

We present plots of the solution u and the computed PML approximation $\widetilde{u}_{t,h}$ for visual comparison. Figure 3.1 presents three-dimensional surface plots of the components of the two functions. As expected, in the domain of interest the functions appear to be identical, while inside the PML layer, the function $\widetilde{u}_{t,h}$ quickly decays to 0. The "cut-off" effect of the PML is clearly seen from the overhead view in Figure 3.2.

In Table 3.1 we have reported the L^2- and H^1- norms of the error inside the domain of interest, as well as the order of convergence. The last column provides the number of iterations needed to obtain convergence. Consistent with Remark 7, it suggests that there is a uniform (in h) bound for the condition number of the preconditioned discrete system.

Table 3.1. H^1- and L^2- errors in the finite element approximation

h	# dofs	L^2-error		H^1-error		# iter
$\frac{1}{2}$	864	0.265284	-	1.43557	-	100
$\frac{1}{4}$	3264	0.156433	0.76	0.707713	1.02	294
$\frac{1}{8}$	12672	0.085375	0.87	0.358187	0.98	840
$\frac{1}{16}$	49920	0.023992	1.83	0.167791	1.09	1503
$\frac{1}{32}$	198144	0.005956	2.01	0.081722	1.03	1794
$\frac{1}{64}$	789504	0.001471	2.02	0.040577	1.01	1830

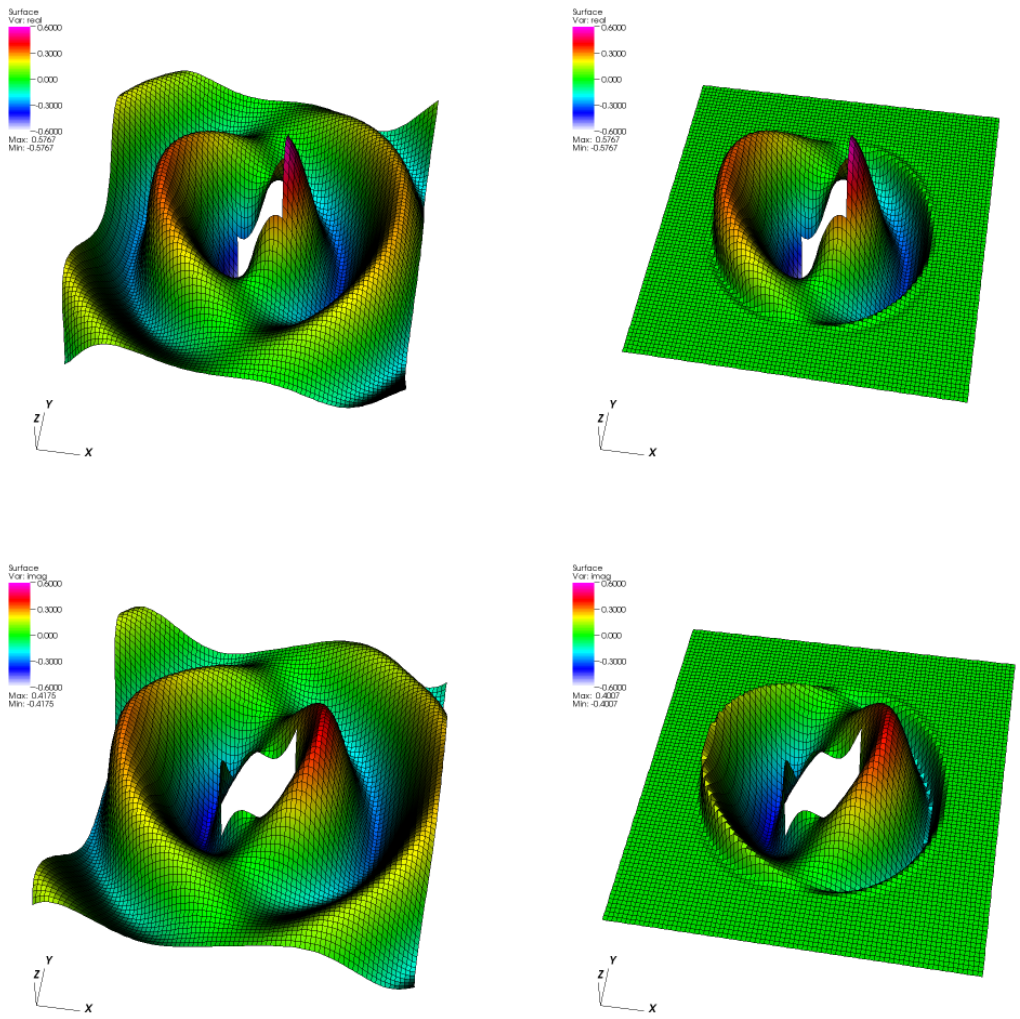

Fig. 3.1. Real (top row) and imaginary (bottom row) components of the solution to the original problem and the PML approximation.

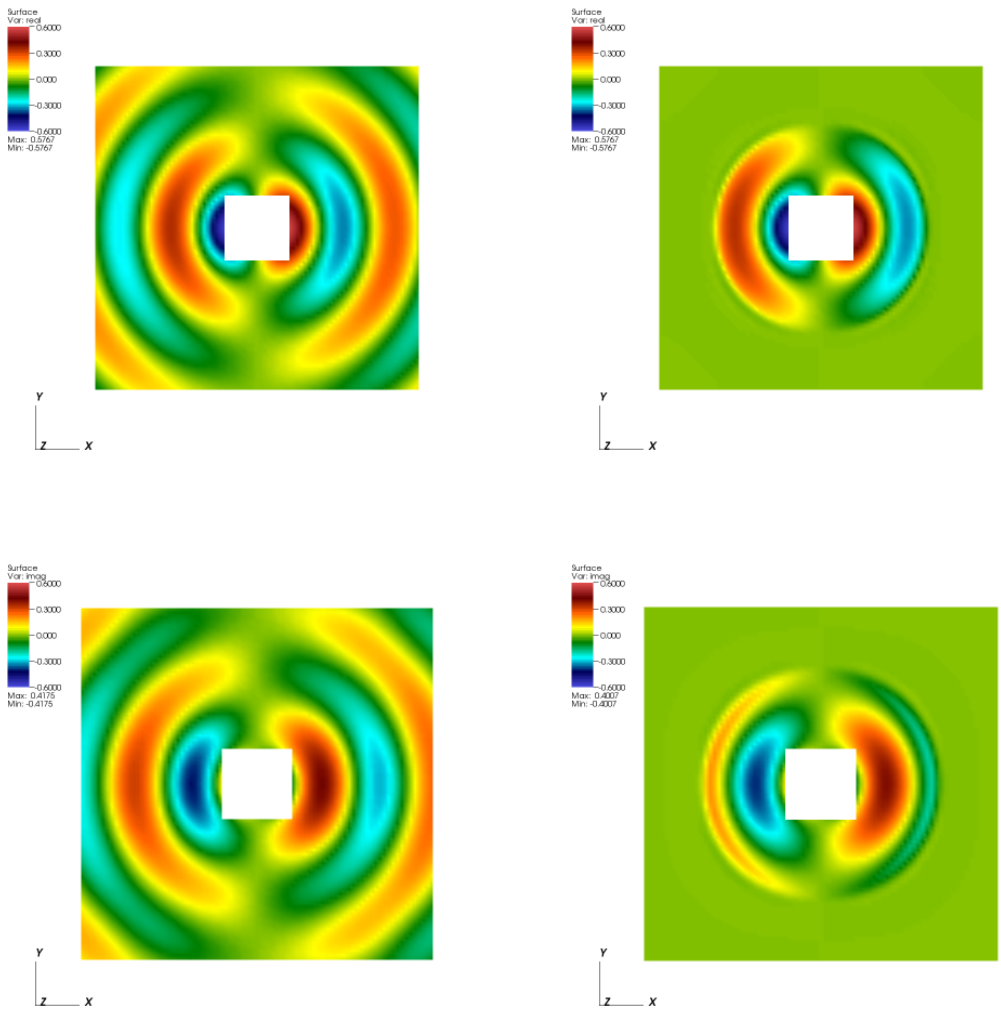

Fig. 3.2. Overhead view of the components of the solution to the original problem and the PML approximation.

CHAPTER IV

ANALYSIS OF THE ELASTIC WAVE PML PROBLEM

In this chapter, we apply the complex shift technique in spherical coordinates to derive and analyze a PML model for the time-harmonic elastic wave problem. Similarly to the Helmholtz PML of the previous chapter, the complex shift technique leads to a single (variable coefficient) equation for the displacement vector and, in contrast to the elastic wave PML models in [12] and [17], requires no field splitting or additional unknowns. A PML model similar to this was proposed in [27] and limited numerical results were reported.

The contents of this chapter and the following one will appear in [8].

4.1. Formulation of the elastic wave problem

In this section, we formulate the elastic wave problem and its far field boundary condition. Let Ω be a bounded domain, with boundary Γ, containing the origin and let Ω^c denote its complement. We seek a vector valued function $\boldsymbol{u} \in \boldsymbol{H}^1_{loc}(\Omega^c)$ satisfying

$$\boldsymbol{\Delta u} + \gamma \boldsymbol{\nabla} \nabla \cdot \boldsymbol{u} + k^2 \boldsymbol{u} = \boldsymbol{0} \text{ in } \Omega^c \tag{4.1}$$

and

$$\boldsymbol{u} = \boldsymbol{g} \text{ on } \Gamma. \tag{4.2}$$

Here γ and k are positive real numbers and $\boldsymbol{g}$ is given in $\boldsymbol{H}^{1/2}(\Gamma)$.

The formulation is completed by imposing the "so-called" Kupradze-Sommerfeld far field radiation condition. This condition involves decomposing the function $\boldsymbol{u}$ away from Ω as

$$\boldsymbol{u} = \boldsymbol{\zeta} + \boldsymbol{\psi} \tag{4.3}$$

with $\boldsymbol{\zeta}$ solenoidal and $\boldsymbol{\psi}$ irrotational. As shown in [7], the components can be chosen to satisfy Helmholtz equations:

$$\boldsymbol{\Delta\zeta} + k^2\boldsymbol{\zeta} = \mathbf{0}$$

and

$$\boldsymbol{\Delta\psi} + k_1^2\boldsymbol{\psi} = \mathbf{0},$$

where $k_1 = k/\sqrt{1+\gamma}$. The Kupradze-Sommerfeld radiation condition imposes the corresponding Sommerfeld radiation conditions on the components, i.e.,

$$\lim_{r\to\infty} r(\frac{\partial\boldsymbol{\zeta}}{\partial n} - ik\boldsymbol{\zeta}) = \mathbf{0} \tag{4.4}$$

and

$$\lim_{r\to\infty} r(\frac{\partial\boldsymbol{\psi}}{\partial n} - ik_1\boldsymbol{\psi}) = \mathbf{0}. \tag{4.5}$$

Let B_R denote the open ball of radius R centered at the origin and assume that $\overline{\Omega}$ is contained in B_R. The above decomposition is, in fact, uniquely determined from the values of $\boldsymbol{u}$ on B_R (see, e.g., [7]).

The components above can be expanded in series outside of B_R. Indeed,

$$\boldsymbol{\psi}(\boldsymbol{x}) = \sum_{n=0}^{\infty}\sum_{|m|\le n} \boldsymbol{\gamma}_{n,m} p_n(r) Y_{n,m}(\hat{\boldsymbol{x}}) \tag{4.6}$$

with $\boldsymbol{\gamma}_{n,m} \in \mathbb{R}^3$. Here $p_n(r) \equiv h_n^{(1)}(k_1 r)$, $h_n^{(1)}$ is the Hankel function of the first kind of order n, $Y_{n,m}$ are spherical harmonics, $r = |\boldsymbol{x}|$ and $\hat{\boldsymbol{x}} = \boldsymbol{x}/r$. Similarly,

$$\boldsymbol{\zeta} = \sum_{n=0}^{\infty}\sum_{|m|\le n} \boldsymbol{\alpha}_{n,m} q_n(r) Y_{n,m}(\hat{\boldsymbol{x}}) \tag{4.7}$$

with $\boldsymbol{\alpha}_{n,m} \in \mathbb{R}^3$ and $q_n(r) \equiv h_n^{(1)}(kr)$. Hence outside of B_R, $\boldsymbol{u}$ may be expanded in a

series of the form

$$\boldsymbol{u} = \sum_{n=0}^{\infty} \sum_{|m|\leq n} (\boldsymbol{\alpha}_{n,m} q_n(r) + \boldsymbol{\gamma}_{n,m} p_n(r)) Y_{n,m}(\hat{\boldsymbol{x}}). \tag{4.8}$$

The following theorem, which will be essential for this chapter, is the main result of [7].

Theorem 5. *For any function* $\boldsymbol{g}$ *in* $\boldsymbol{H}^{1/2}(\Gamma)$*, there is a unique solution* $\boldsymbol{u}$ *in* $\boldsymbol{H}^1_{loc}(\Omega^c)$ *to the elastic wave problem* (4.1)*, satisfying conditions* (4.2)*,* (4.4) *and* (4.5)*. Moreover, for any* R*,* $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega^c \cap B_R)$ *and satisfies*

$$\|\boldsymbol{u}\|_{1,(\Omega^c \cap B_R)} \leq C(R) \|\boldsymbol{g}\|_{1/2,\Gamma}. \tag{4.9}$$

4.2. The elastic wave PML problem

Throughout this chapter, we shall use a sequence of finite subdomains of Ω^c with spherical outer boundaries. Let $r_{-1} < r_0 < r_1 < r_2$ be an increasing sequence of positive real numbers and let Ω_i denote (the interior of) the open ball B_i of radius r_i excluding $\overline{\Omega}$ (we assume that r_{-1} is large enough so that the corresponding ball contains $\overline{\Omega}$). We denote the outer boundary of Ω_i by Γ_i. The values of r_{-1}, r_0, r_1 are independent of the computational outer boundary scaling parameter R_t (introduced in Section 4.3).

As in the case of the Helmholtz equation, the differential operators involved in the PML approximations are defined in terms of a formal complex change of variables, based on the transformation (2.1) with $s = i$. We will, once again, be considering a spherical PML, taking $n(\boldsymbol{x}) = \|\boldsymbol{x}\|_2 = r$ and $\widetilde{\sigma}(r) = \widetilde{\sigma}_{pml}(r - r_0)$.

Recall the notation

$$\boldsymbol{A} = \frac{1}{\widetilde{d}^2} \boldsymbol{P} + \frac{1}{d\widetilde{d}} (\boldsymbol{I} - \boldsymbol{P}),$$
$$\boldsymbol{B} = d\boldsymbol{P}\boldsymbol{u} + \widetilde{d}(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{u},$$

where $\boldsymbol{P}$, given in (2.10), was the matrix diagonal in spherical coordinates with value one on the radial component and zero on the angular ones.

In addition to the shifted gradient and divergence operators, given in (2.12), we shall also need the shifted versions of the vector Laplacian and the curl operators given below

$$\widetilde{\boldsymbol{\Delta}}\boldsymbol{w} = \begin{pmatrix} \widetilde{\Delta} w_1 \\ \widetilde{\Delta} w_2 \\ \widetilde{\Delta} w_3 \end{pmatrix} \text{ and } \widetilde{\boldsymbol{\nabla}} \times \boldsymbol{F} = \boldsymbol{A}\, \boldsymbol{\nabla} \times (\boldsymbol{B}\boldsymbol{F}).$$

As in Theorem 3, one can verify the identity

$$\widetilde{\boldsymbol{\Delta}}\widetilde{\boldsymbol{u}} + \gamma \widetilde{\boldsymbol{\nabla}} \widetilde{\nabla} \cdot \widetilde{\boldsymbol{u}} + k^2 \widetilde{\boldsymbol{u}} = \boldsymbol{0} \quad \text{ in } \Omega^c, \tag{4.10}$$

for the "stretched" function $\widetilde{\boldsymbol{u}}$, corresponding to a function $\boldsymbol{u}$ satisfying (4.1), defined by

$$\widetilde{\boldsymbol{u}}(\boldsymbol{x}) = \begin{cases} \boldsymbol{u}(\boldsymbol{x}): & \text{if } \boldsymbol{x} \in \Omega^c \text{ and } |\boldsymbol{x}| \leq r_0, \\ \displaystyle\sum_{n=0}^{\infty} \sum_{|m| \leq n} (\boldsymbol{\alpha}_{n,m} q_n(\tilde{r}) + \boldsymbol{\gamma}_{n,m} p_n(\tilde{r})) Y_{n,m}(\hat{\boldsymbol{x}}): & \text{otherwise.} \end{cases} \tag{4.11}$$

Here $\tilde{r} = (1 + i\widetilde{\sigma}(r))r$ and $\{\boldsymbol{\gamma}_{n,m}\}$ and $\{\boldsymbol{\alpha}_{n,m}\}$ are the coefficients appearing in (4.6) and (4.7) respectively.

The analyticity arguments in the proof of Theorem 3 can also be used to verify the following three identities for the stretched operators, which we shall require for our analysis:

$$\widetilde{\boldsymbol{\Delta}}\boldsymbol{w} = -\widetilde{\boldsymbol{\nabla}} \times \widetilde{\boldsymbol{\nabla}} \times \boldsymbol{w} + \widetilde{\boldsymbol{\nabla}} \widetilde{\nabla} \cdot \boldsymbol{w}, \tag{4.12}$$

$$((\det \boldsymbol{J}) \widetilde{\boldsymbol{\Delta}}\boldsymbol{w}, \widetilde{\boldsymbol{\nabla}}\phi) = ((\det \boldsymbol{J}) \widetilde{\boldsymbol{\nabla}} (\widetilde{\nabla} \cdot \boldsymbol{w}), \widetilde{\boldsymbol{\nabla}}\phi), \tag{4.13}$$

and

$$((\det \boldsymbol{J})\widetilde{\boldsymbol{\Delta}}\boldsymbol{w}, \widetilde{\boldsymbol{\nabla}} \times \boldsymbol{\phi}) = -\sum_i ((\det \boldsymbol{J})\widetilde{\boldsymbol{\nabla}}(\widetilde{\boldsymbol{\nabla}} \times \boldsymbol{w})_i, \widetilde{\boldsymbol{\nabla}}\boldsymbol{\phi}_i). \tag{4.14}$$

Here $\boldsymbol{w} \in \boldsymbol{H}^2_{loc}(\mathbb{R}^3)$ and $\boldsymbol{\phi}$ and ϕ are smooth and compactly supported. Identities (4.13) and (4.14) basically state that the scaled Laplacian (viewed as an operator on both scalar and vector functions) commutes with the scaled curl and divergence operators. We have stated them in weak form (in the appropriate inner product) as not to require additional smoothness on $\boldsymbol{w}$. Since we are using spherical scaling, $\det \boldsymbol{J} = \widetilde{d}^2 d$ (or, in 2D, $\widetilde{d}d$).

For $\boldsymbol{v}, \boldsymbol{\Psi} \in \boldsymbol{H}^1(\mathbb{R}^3)$, we define

$$\begin{aligned}\mathcal{A}(\boldsymbol{v}, \boldsymbol{\Psi}) \equiv &\sum_{j=1}^{3} (\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}\boldsymbol{\nabla}\,\boldsymbol{v}_j, \boldsymbol{\nabla}\,\boldsymbol{\Psi}_j)_{\mathbb{R}^3} \\ &+ \gamma((\tilde{d}^2 d)^{-1}\nabla \cdot (\boldsymbol{A}^{-1}\boldsymbol{v}), \nabla \cdot (\overline{\boldsymbol{A}}^{-1}\boldsymbol{\Psi}))_{\mathbb{R}^3} - k^2(\widetilde{d}^2 d\boldsymbol{v}, \boldsymbol{\Psi})_{\mathbb{R}^3}.\end{aligned} \tag{4.15}$$

Here $\overline{\boldsymbol{A}}^{-1}$ denotes the complex conjugate. Note that all terms in the above expressions are non-Hermitian.

We shall use the form $\mathcal{A}(\cdot, \cdot)$ to denote the restriction of the above integrals to $D \subseteq \mathbb{R}^3$ when one of the two arguments is in $\boldsymbol{H}^1_0(D)$ and the other is in $\boldsymbol{H}^1(D)$.

Multiplying (4.10) by $\tilde{d}^2 d\bar{\boldsymbol{\Psi}}$ for $\boldsymbol{\Psi} \in \boldsymbol{C}^\infty_0(\Omega^c)$ and integrating by parts shows that $\widetilde{\boldsymbol{u}}$ satisfies the weak equations

$$\mathcal{A}(\widetilde{\boldsymbol{u}}, \boldsymbol{\Psi}) = 0.$$

This equality, of course, extends to all $\boldsymbol{\Psi} \in \boldsymbol{H}^1_0(D)$ by density.

For the first step in our analysis we consider the source problem on all of $\mathbb{R}^3$,

$$\widetilde{\boldsymbol{\Delta}}\boldsymbol{U} + \gamma\widetilde{\boldsymbol{\nabla}}\widetilde{\nabla} \cdot \boldsymbol{U} + k^2\boldsymbol{U} = \boldsymbol{\Phi} \quad \text{ in } \mathbb{R}^3. \tag{4.16}$$

A weak form of this equation is

$$\mathcal{A}(\boldsymbol{U}, \boldsymbol{\Psi}) = (\tilde{d}^2 d\boldsymbol{\Phi}, \boldsymbol{\Psi})_{\mathbb{R}^3}, \quad \text{for all } \boldsymbol{\Psi} \in \boldsymbol{H}^1(\mathbb{R}^3). \tag{4.17}$$

Let $\widetilde{H}^1(\mathbb{R}^3)$ denote the weighted Sobolev space of functions defined on $\mathbb{R}^3$ given by

$$\widetilde{H}^1(\mathbb{R}^3) = \{u : \ u(1+r^2)^{-1/2} \in L^2(\mathbb{R}^3) \text{ and } \boldsymbol{\nabla} u \in \boldsymbol{L}^2(\mathbb{R}^3)\}.$$

It follows from Theorem 2.5.13 of [20] that $\|\boldsymbol{\nabla}\,\phi\|_{\boldsymbol{L}^2(\mathbb{R}^3)}$ provides an equivalent norm for $\widetilde{H}^1(\mathbb{R}^3)$. Moreover, $C_0^\infty(\mathbb{R}^3)$ is dense in $\widetilde{H}^1(\mathbb{R}^3)$ [20].

For the analysis of the above problem, we shall need to decompose a vector function $\boldsymbol{\Phi} \in \boldsymbol{L}^2(\mathbb{R}^3)$ as

$$\boldsymbol{\Phi} = \widetilde{\boldsymbol{\nabla}}\theta + \boldsymbol{\Phi}_0 \tag{4.18}$$

where $\theta \in \widetilde{H}^1(\mathbb{R}^3)$. We choose θ so that $\widetilde{\nabla} \cdot \boldsymbol{\Phi}_0 = 0$, specifically, $\theta \in \widetilde{H}^1(\mathbb{R}^3)$ is the solution of

$$\tilde{b}(\theta, \phi) = (\boldsymbol{A}^{-1}\boldsymbol{\Phi}, \boldsymbol{\nabla}\,\phi)_{\mathbb{R}^3}, \quad \text{for all } \phi \in \widetilde{H}^1(\mathbb{R}^3). \tag{4.19}$$

Here

$$\tilde{b}(\theta, \phi) \equiv (\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\nabla}\,\theta, \boldsymbol{\nabla}\,\phi)_{\mathbb{R}^3}.$$

Since $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ has a uniformly positive definite real part, $\tilde{b}$ is coercive on $\widetilde{H}^1(\mathbb{R}^3)$. It follows that (4.19) has a unique solution $\theta \in \widetilde{H}^1(\mathbb{R}^3)$ satisfying

$$\|\boldsymbol{\nabla}\,\theta\|_{\boldsymbol{L}^2(\mathbb{R}^3)} \leq C\|\boldsymbol{\Phi}\|_{\boldsymbol{L}^2(\mathbb{R}^3)}. \tag{4.20}$$

We shall require the following lemma whose proof we provide later.

Lemma 2. *The decomposition* (4.18) *is stable in* $\boldsymbol{H}^{-1}(\mathbb{R}^3)$*; i.e.*

$$\|\boldsymbol{\Phi}_0\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} + \|\widetilde{\boldsymbol{\nabla}}\theta\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} \leq C\|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)}.$$

Using the above lemma, we can prove the following theorem.

Theorem 6. *Let* $\boldsymbol{\Phi} \in \boldsymbol{H}^{-1}(\mathbb{R}^3)$. *Problem* (4.17) *has a solution* $\boldsymbol{U} \in \boldsymbol{H}^1(\mathbb{R}^3)$ *satisfying*

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C\|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)}. \tag{4.21}$$

Proof. In Section 3.3 we analyzed the 2D scalar Helmholtz equation. As stated in Remark 4, the results there extend to the 3D equation

$$\widetilde{\Delta} u + k^2 u = \Phi, \text{ in } \mathbb{R}^3. \tag{4.22}$$

and its variational formulation: Find $u \in H^1(\mathbb{R}^3)$ satisfying

$$b(u, \theta) = (\tilde{d}^2 \Phi, \theta)_{\mathbb{R}^3}, \quad \text{for all } \theta \in H^1(\mathbb{R}^3). \tag{4.23}$$

Here the form $b(\cdot, \cdot)$ is the one given in (3.29)

$$b(u, \theta) = (\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}\boldsymbol{\nabla}\, u, \boldsymbol{\nabla}\, (\bar{d}^{-1}\theta))_{\mathbb{R}^3} - k^2(\widetilde{d}^2 u, \theta)_{\mathbb{R}^3}.$$

We shall use in particular, that the solution u of the above problem is unique, belongs to $H^2(\mathbb{R}^3)$ and satisfies

$$\|u\|_{H^2(\mathbb{R}^3)} \leq C\|\Phi\|_{L^2(\mathbb{R}^3)} \tag{4.24}$$

(see Theorem 4).

We first construct a solution of (4.17). We decompose $\boldsymbol{\Phi} \in L^2(\mathbb{R}^3)$ as follows:

$$\boldsymbol{\Phi} = \widetilde{\boldsymbol{\nabla}}\theta + (\boldsymbol{\Phi} - \widetilde{\boldsymbol{\nabla}}\theta) \equiv \widetilde{\boldsymbol{\nabla}}\theta + \boldsymbol{\Phi}_0$$

where θ is the solution of (4.19). By construction, $\widetilde{\nabla} \cdot \boldsymbol{\Phi}_0 = 0$.

We then define $\boldsymbol{U}_1$ by

$$\widetilde{\boldsymbol{\Delta}}\boldsymbol{U}_1 + k^2\boldsymbol{U}_1 = \boldsymbol{\Phi}_0. \tag{4.25}$$

This equation reduces component-wise to the scalar Helmholtz PML equation with Φ replaced by $(\boldsymbol{\Phi}_0)_j$, $j = 1, 2, 3$. Thus $\boldsymbol{U}_1 \in \boldsymbol{H}^2(\mathbb{R}^3)$ and

$$\|\boldsymbol{U}_1\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C\|\boldsymbol{\Phi}_0\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)}. \tag{4.26}$$

Multiplying (4.25) by $\widetilde{d}^2 d\widetilde{\boldsymbol{\nabla}}(d^{-1}\bar{\phi})$, integrating and applying (4.13) gives that $\widetilde{\nabla} \cdot \boldsymbol{U}_1$ satisfies

$$b(\widetilde{\nabla} \cdot \boldsymbol{U}_1, \phi) = 0 \quad \text{for all } \phi \in H^1(\mathbb{R}^3).$$

Uniqueness implies that $\widetilde{\nabla} \cdot \boldsymbol{U}_1 = 0$ so $\boldsymbol{U}_1$ satisfies

$$\widetilde{\boldsymbol{\Delta}}\boldsymbol{U}_1 + \gamma\widetilde{\boldsymbol{\nabla}}(\widetilde{\nabla} \cdot \boldsymbol{U}_1) + k^2\boldsymbol{U}_1 = \boldsymbol{\Phi}_0 \text{ in } \mathbb{R}^3. \tag{4.27}$$

Next define $\boldsymbol{U}_2$ to be the solution of

$$(1+\gamma)\widetilde{\boldsymbol{\Delta}}\boldsymbol{U}_2 + k^2\boldsymbol{U}_2 = \widetilde{\boldsymbol{\nabla}}\theta. \tag{4.28}$$

This equation also reduces component-wise to (4.22) with k^2 replaced by $k^2(1+\gamma)^{-1}$ and Φ replaced by $(1+\gamma)^{-1}(\widetilde{\boldsymbol{\nabla}}\theta)_i$, $i = 1, 2, 3$. Thus $\boldsymbol{U}_2 \in \boldsymbol{H}^2(\mathbb{R}^3)$ and

$$\|\boldsymbol{U}_2\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C\|\widetilde{\boldsymbol{\nabla}}\theta\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)}. \tag{4.29}$$

Multiplying (4.28) by $\widetilde{d}^2 d\widetilde{\boldsymbol{\nabla}} \times (d^{-1}\bar{\boldsymbol{\phi}})$, integrating and applying (4.14) shows that each component of $\widetilde{\boldsymbol{\nabla}} \times \boldsymbol{U}_2$ satisfies

$$b((\widetilde{\boldsymbol{\nabla}} \times \boldsymbol{U}_2)_j, \phi) = 0 \quad \text{for all } \phi \in H^1(\mathbb{R}^3).$$

with b defined using k_1. As above, this implies that $\widetilde{\boldsymbol{\nabla}} \times \boldsymbol{U}_2 = \boldsymbol{0}$. It follows from

(4.28)) and (4.12) that $\boldsymbol{U}_2$ satisfies

$$\widetilde{\boldsymbol{\Delta}}\boldsymbol{U}_2 + \gamma\widetilde{\boldsymbol{\nabla}}(\widetilde{\nabla}\cdot\boldsymbol{U}_2) + k^2\boldsymbol{U}_2 = \widetilde{\boldsymbol{\nabla}}\theta \text{ in } \mathbb{R}^3.$$

Finally, we define $\boldsymbol{U} = \boldsymbol{U}_1 + \boldsymbol{U}_2$. Clearly, $\boldsymbol{U}$ satisfies (4.16) and (4.21) follows from (4.26), (4.29) and Lemma 2. This completes the proof of the theorem. □

Proof of Lemma 2. It suffices to show that

$$\|\boldsymbol{\Phi}_0\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} \leq C\|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)}. \tag{4.30}$$

Given $\boldsymbol{\Psi} \in \boldsymbol{H}^1(\mathbb{R}^3)$, let $\eta \in \widetilde{H}^1(\mathbb{R}^3)$ solve the (adjoint) problem

$$\tilde{b}(\delta,\eta) = (\widetilde{\boldsymbol{\nabla}}\delta, \boldsymbol{\Psi})_{\mathbb{R}^3}, \quad \text{for all } \delta \in \widetilde{H}^1(\mathbb{R}^3).$$

Clearly,

$$\|\boldsymbol{\nabla}\,\eta\|_{\boldsymbol{L}^2(\mathbb{R}^3)} \leq C\|\boldsymbol{\Psi}\|_{\boldsymbol{L}^2(\mathbb{R}^3)}. \tag{4.31}$$

Moreover, for any first difference quotient D_h of size $h < 1$, it follows easily from integration by parts and (4.31) that

$$|\tilde{b}(\delta, D_h\eta)| \leq C\|\boldsymbol{\nabla}\,\delta\|_{\boldsymbol{L}^2(\mathbb{R}^3)}\|\boldsymbol{\Psi}\|_{\boldsymbol{H}^1(\mathbb{R}^3)},$$

with C independent of h. From this it follows that

$$\|\boldsymbol{\nabla}\,\eta\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C\|\boldsymbol{\Psi}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}. \tag{4.32}$$

We then have

$$\begin{aligned}\|\boldsymbol{\Phi}_0\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} &= \sup_{\boldsymbol{\Psi}\in\boldsymbol{H}^1(\mathbb{R}^3)} \frac{|(\boldsymbol{\Phi}_0,\boldsymbol{\Psi})_{\mathbb{R}^3}|}{\|\boldsymbol{\Psi}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}} = \sup_{\boldsymbol{\Psi}\in\boldsymbol{H}^1(\mathbb{R}^3)} \frac{|(\boldsymbol{\Phi}_0,\boldsymbol{\Psi}-\bar{A}^{-1}\boldsymbol{\nabla}\,\eta)_{\mathbb{R}^3}|}{\|\boldsymbol{\Psi}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}} \\ &= \sup_{\boldsymbol{\Psi}\in\boldsymbol{H}^1(\mathbb{R}^3)} \frac{|(\boldsymbol{\Phi},\boldsymbol{\Psi}-\bar{A}^{-1}\boldsymbol{\nabla}\,\eta)_{\mathbb{R}^3}|}{\|\boldsymbol{\Psi}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}}.\end{aligned}$$

Combining this with (4.32) gives (4.30) and completes the proof of the lemma. □

We have constructed a solution of (4.17) which satisfies (4.21). We will show that this is the unique solution of (4.17).

Theorem 7. *Let* $\boldsymbol{W} \in \boldsymbol{H}^1(\mathbb{R}^3)$ *satisfy* $\mathcal{A}(\boldsymbol{W}, \boldsymbol{\Theta}) = 0$ *for all* $\boldsymbol{\Theta} \in \boldsymbol{H}^1(\mathbb{R}^3)$. *Then* $\boldsymbol{W} = \boldsymbol{0}$.

Proof. Let $\boldsymbol{W}$ be as above. By Theorem 6, there is a solution $\boldsymbol{x}$ of

$$\mathcal{A}(\boldsymbol{x}, \boldsymbol{\Theta}) = (\bar{\boldsymbol{W}}, \boldsymbol{\Theta})_{\mathbb{R}^3} \quad \text{for all } \boldsymbol{\Theta} \in \boldsymbol{H}^1(\mathbb{R}^3).$$

Thus

$$\|\bar{\boldsymbol{W}}\|^2_{\boldsymbol{L}^2(\mathbb{R}^3)} = \mathcal{A}(\boldsymbol{W}, \bar{\boldsymbol{x}}) = 0.$$

Hence $\boldsymbol{W} = \boldsymbol{0}$. □

It will be useful to restate the last two results as follows.

Theorem 8. *Let* $\boldsymbol{U}$ *be in* $\boldsymbol{H}^1(\mathbb{R}^3)$. *Then*

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C \sup_{\boldsymbol{V} \in \boldsymbol{H}^1(\mathbb{R}^3)} \frac{|\mathcal{A}(\boldsymbol{U}, \boldsymbol{V})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}} \tag{4.33}$$

and

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C \sup_{\boldsymbol{V} \in \boldsymbol{H}^1(\mathbb{R}^3)} \frac{|\mathcal{A}(\boldsymbol{V}, \boldsymbol{U})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}}. \tag{4.34}$$

Proof. Let $\boldsymbol{U}$ be in $\boldsymbol{H}^1(\mathbb{R}^3)$. Multiplication by a uniformly bounded bijective C^1 function whose inverse is also uniformly bounded is an isomorphism of $H^1(\mathbb{R}^3)$ onto $H^1(\mathbb{R}^3)$ and hence is also an isomorphism of $H^{-1}(\mathbb{R}^3)$ onto $H^{-1}(\mathbb{R}^3)$. Thus, $\mathcal{A}(\boldsymbol{U}, \cdot)$ defines an element $\boldsymbol{\Phi}$ of $\boldsymbol{H}^{-1}(\mathbb{R}^3)$ by

$$\langle \widetilde{d}^2 d\boldsymbol{\Phi}, \boldsymbol{V} \rangle = \mathcal{A}(\boldsymbol{U}, \boldsymbol{V}).$$

Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing. The above corollary shows that $\boldsymbol{U}$ coincides with the function constructed in Theorem 6. Thus,

$$\begin{aligned}\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} &\leq C\|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} \leq C\|\widetilde{d^2}d\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} \\ &= C \sup_{\boldsymbol{V}\in\boldsymbol{H}^1(\mathbb{R}^3)} \frac{|\mathcal{A}(\boldsymbol{U},\boldsymbol{V})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}}.\end{aligned}$$

This is the first *inf-sup* condition.

The second *inf-sup* condition follows from the first, indeed,

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} = \|\bar{\boldsymbol{U}}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C \sup_{\boldsymbol{V}\in\boldsymbol{H}^1(\mathbb{R}^3)} \frac{|\mathcal{A}(\bar{\boldsymbol{U}},\boldsymbol{V})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}} = \sup_{\boldsymbol{V}\in\boldsymbol{H}^1(\mathbb{R}^3)} \frac{|\mathcal{A}(\boldsymbol{V},\boldsymbol{U})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}}.$$

□

We next consider the form (4.15) restricted to $\boldsymbol{H}_0^1(\Omega^c)$ and prove the analogous theorem.

Theorem 9. *Let $\boldsymbol{U}$ be in $\boldsymbol{H}_0^1(\Omega^c)$. Then*

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega^c)} \leq C \sup_{\boldsymbol{V}\in\boldsymbol{H}_0^1(\Omega^c)} \frac{|\mathcal{A}(\boldsymbol{U},\boldsymbol{V})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega^c)}} \tag{4.35}$$

and

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega^c)} \leq C \sup_{\boldsymbol{V}\in\boldsymbol{H}_0^1(\Omega^c)} \frac{|\mathcal{A}(\boldsymbol{V},\boldsymbol{U})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega^c)}}. \tag{4.36}$$

Proof. Let $\boldsymbol{\Phi} \in \boldsymbol{H}^{-1}(\Omega^c)$. By the Hahn-Banach Theorem we may extend $\boldsymbol{\Phi}$ to $\boldsymbol{H}^{-1}(\mathbb{R}^3)$ with

$$\|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\mathbb{R}^3)} = \|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\Omega^c)}.$$

By Theorem 6, there is a unique solution $\boldsymbol{W}$ of

$$\mathcal{A}(\boldsymbol{W},\boldsymbol{\Theta}) = \langle \boldsymbol{\Phi}, \boldsymbol{\Theta} \rangle, \quad \text{for all } \boldsymbol{\Theta} \in \boldsymbol{H}^1(\mathbb{R}^3).$$

Now let $\boldsymbol{U}_0$ be the solution of the elastic wave problem satisfying (4.1), (4.4), (4.5) and (4.2) with $\boldsymbol{g} = \boldsymbol{W}$ on Γ. Also let $\widetilde{\boldsymbol{U}_0}$ be the solution of the corresponding PML

elastic wave problem given by (4.11) (with $\boldsymbol{u}$ replaced by $\boldsymbol{U}_0$). By (4.9),

$$\|\widetilde{\boldsymbol{U}_0}\|_{\boldsymbol{H}^1(\Omega_0)} = \|\boldsymbol{U}_0\|_{\boldsymbol{H}^1(\Omega_0)} \leq C\|\boldsymbol{W}\|_{\boldsymbol{H}^{1/2}(\Gamma)}. \tag{4.37}$$

We will show that

$$\|\widetilde{\boldsymbol{U}_0}\|_{\boldsymbol{H}^1(\Omega^c)} \leq C\|\boldsymbol{W}\|_{\boldsymbol{H}^{1/2}(\Gamma)}. \tag{4.38}$$

Let ψ be a smooth cut-off function satisfying $\psi(\boldsymbol{x}) = 1$ for $|\boldsymbol{x}| \geq r_0$ and $\psi(\boldsymbol{x}) = 0$ on Γ. Let $\boldsymbol{V} \equiv \psi\widetilde{\boldsymbol{U}}_0$ in Ω^c be extended by $\boldsymbol{0}$ to all of $\mathbb{R}^3$. Let $\boldsymbol{\phi}$ be in $\boldsymbol{H}^1(\mathbb{R}^3)$. Since

$$\mathcal{A}(\widetilde{\boldsymbol{U}_0}, \boldsymbol{\Theta}) = 0, \quad \text{for all } \boldsymbol{\Theta} \in \boldsymbol{H}_0^1(\Omega^c)$$

it follows that

$$|\mathcal{A}(\boldsymbol{V}, \boldsymbol{\phi})| = |\mathcal{A}((\psi - 1)\widetilde{\boldsymbol{U}}_0, \boldsymbol{\phi}) + \mathcal{A}(\widetilde{\boldsymbol{U}}_0, (1 - \psi)\boldsymbol{\phi})| \leq C\|\widetilde{\boldsymbol{U}}_0\|_{\boldsymbol{H}^1(\Omega_0)}\|\boldsymbol{\phi}\|_{\boldsymbol{H}^1(\mathbb{R}^3)}.$$

Applying (4.33) gives

$$\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\mathbb{R}^3)} \leq C\|\widetilde{\boldsymbol{U}}_0\|_{\boldsymbol{H}^1(\Omega_0)}.$$

We then have

$$\|\widetilde{\boldsymbol{U}}_0\|_{\boldsymbol{H}^1(\Omega^c)} \leq \|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega^c)} + \|(1 - \psi)\widetilde{\boldsymbol{U}}_0\|_{\boldsymbol{H}^1(\Omega_0)} \leq C\|\widetilde{\boldsymbol{U}}_0\|_{\boldsymbol{H}^1(\Omega_0)}.$$

and (4.38) follows from (4.37).

Now set $\boldsymbol{U} = \boldsymbol{W} - \widetilde{\boldsymbol{U}_0}$. We have constructed a vector function $\boldsymbol{U} \in \boldsymbol{H}^1(\Omega^c)$ such that

$$\mathcal{A}(\boldsymbol{U}, \boldsymbol{\Theta}) = \langle \boldsymbol{\Phi}, \boldsymbol{\Theta} \rangle, \quad \text{for all } \boldsymbol{\Theta} \in \boldsymbol{H}_0^1(\Omega^c),$$

for any $\boldsymbol{\Phi} \in \boldsymbol{H}^{-1}(\Omega^c)$. Also

$$\begin{aligned} \|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega^c)} &\leq \|\boldsymbol{W}\|_{\boldsymbol{H}^1(\Omega^c)} + \|\widetilde{\boldsymbol{U}_0}\|_{\boldsymbol{H}^1(\Omega^c)} \\ &\leq C(\|\boldsymbol{W}\|_{\boldsymbol{H}^1(\Omega^c)} + \|\boldsymbol{W}\|_{\boldsymbol{H}^{1/2}(\Gamma)}) \leq C\|\boldsymbol{\Phi}\|_{\boldsymbol{H}^{-1}(\Omega^c)}. \end{aligned} \tag{4.39}$$

The arguments proving Theorems 7 and 8 now complete the proof of the theorem. □

4.3. The truncated PML problem

In this section, we consider approximating the PML problem on Ω^c by a problem on a truncated domain with a convenient boundary condition on the outer boundary. For convenience, we shall use homogeneous Dirichlet conditions. Let Ω_∞ be a bounded subset of Ω^c containing the transition region, i.e., $\Omega_1 \subseteq \Omega_\infty$. The size of Ω_∞ is controlled by a parameter R_t, for example, the outer boundary Γ_∞ of Ω_∞ could be a cube with edge of length $2R_t$. In any event, we assume that Γ_∞ is uniformly Lipschitz with constants that are independent of R_t and that $B_{R_t} \subseteq \Omega_\infty$.

We are interested in studying the truncated variational problem, find $\boldsymbol{U} \in \boldsymbol{H}_0^1(\Omega_\infty)$ satisfying

$$\mathcal{A}(\boldsymbol{U}, \boldsymbol{\phi}) = \langle \boldsymbol{F}, \boldsymbol{\phi} \rangle, \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega_\infty). \tag{4.40}$$

One of our main tasks will be to show that this problem is well posed for R_t sufficiently large. This will be a consequence of a result analogous to Theorem 8 and Theorem 9 for the domain Ω_∞. To that end, we shall need the following decay estimate associated with elastic wave equation.

Lemma 3. *Assume that $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega_1^c)$ and satisfies*

$$\boldsymbol{\Delta u} + \gamma \boldsymbol{\nabla} \nabla \cdot \boldsymbol{u} + k^2 d_0^2 \boldsymbol{u} = \boldsymbol{0} \tag{4.41}$$

in Ω_1^c. Then,

$$\|\boldsymbol{u}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)} \leq C e^{-\sigma_0 k_1 R_t} \|\boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega_\infty)}.$$

In order to prove Lemma 3, we will need an interior estimate result for the

solutions to the elastic wave equation, which we state here. It follows immediately from Proposition 3 and the decomposition (4.3).

Proposition 5. *Suppose that* $\boldsymbol{v}$ *satisfies the equation*

$$\boldsymbol{\Delta v} + \gamma \boldsymbol{\nabla} \nabla \cdot \boldsymbol{v} + \beta \boldsymbol{v} = \boldsymbol{0} \tag{4.42}$$

in a domain D *with a (possibly complex) constant* β*. If* D_1 *is a subdomain, whose closure is contained in* D*, then*

$$\|\boldsymbol{v}\|_{H^1(D_1)} \leq C \|\boldsymbol{v}\|_{L^2(D)}. \tag{4.43}$$

Proof of Lemma 3. In the proof, we will use additional expanding domains Ω_i, $i = 3, \ldots, 6$ corresponding to $r_2 < r_3 \cdots < r_6 < R_t$. We use the decomposition given in (4.3) and will prove the result, by using the decay estimate for the Helmholtz PML equation. We will be applying Proposition 4 with $a = r_3$, $b = r_5$ and $D = \Omega_\infty \cup \Omega$ (note that $\partial D = \Gamma_\infty$). Recall the notation $S_\alpha = \{\boldsymbol{x} \ : \ \text{dist}(\boldsymbol{x}, \Gamma_\infty) < \alpha\}$. The value of α will be taken small enough so that $\overline{S}_\alpha$ is in $\Omega^c \setminus \overline{\Omega}_6$.

As seen in [7], $\phi \equiv k_1^{-2} \nabla \cdot \boldsymbol{u}$ is a scalar potential for the irrotational part $\boldsymbol{\psi}$, and it, along with each Cartesian component of $\boldsymbol{\psi} = \boldsymbol{\nabla} \phi$, satisfies (3.31) with $\beta = k_1^2 d_0^2$. Hence by Proposition 3, Proposition 4 and a trace inequality, we see that

$$\begin{aligned} \|\boldsymbol{\psi}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)} &\leq C \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(S_{\alpha/2})} \leq C \|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(S_\alpha)} \leq C e^{-\sigma_0 k_1 R_t} \|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(\Omega_5 \setminus \overline{\Omega}_3)} \\ &\leq C e^{-\sigma_0 k_1 R_t} \|\phi\|_{H^1(\Omega_5 \setminus \overline{\Omega}_3)} \leq C e^{-\sigma_0 k_1 R_t} \|\phi\|_{L^2(\Omega_6 \setminus \overline{\Omega}_2)} \\ &\leq C e^{-\sigma_0 k_1 R_t} \|\nabla \cdot \boldsymbol{u}\|_{L^2(\Omega_6 \setminus \overline{\Omega}_2)} \leq C e^{-\sigma_0 k_1 R_t} \|\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega_\infty)}. \end{aligned} \tag{4.44}$$

For the last inequality we used Proposition 5, since $\boldsymbol{u}$ satisfies (4.42) with $\beta = k^2 d_0^2$. Similarly, each Cartesian component of $\boldsymbol{\zeta}$ satisfies (3.31) with $\beta = k^2 d_0^2$, so that by

Proposition 3, Proposition 4 and a trace inequality, we have

$$\begin{aligned}\|\boldsymbol{\zeta}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)} &\le C\|\boldsymbol{\zeta}\|_{\boldsymbol{H}^1(S_{\alpha/2})} \le C\|\boldsymbol{\zeta}\|_{\boldsymbol{L}^2(S_\alpha)} \le Ce^{-\sigma_0 kR_t}\|\boldsymbol{\zeta}\|_{\boldsymbol{L}^2(\Omega_5\setminus\overline{\Omega}_3)} \\ &\le Ce^{-\sigma_0 kR_t}(\|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(\Omega_5\setminus\overline{\Omega}_3)} + \|\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega_5)}) \le Ce^{-\sigma_0 kR_t}\|\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega_\infty)}.\end{aligned} \tag{4.45}$$

Combining (4.44) and (4.45), and noting that $k_1 < k$, completes the proof of Lemma 3.

□

We are now ready to prove that the variational problem (4.40) is well posed.

Theorem 10. *Let $\boldsymbol{U}$ be in $\boldsymbol{H}^1_0(\Omega_\infty)$. Then, for R_t sufficiently large,*

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega_\infty)} \le C \sup_{\boldsymbol{V}\in\boldsymbol{H}^1_0(\Omega_\infty)} \frac{|\mathcal{A}(\boldsymbol{U},\boldsymbol{V})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega_\infty)}} \tag{4.46}$$

and

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega_\infty)} \le C \sup_{\boldsymbol{V}\in\boldsymbol{H}^1_0(\Omega_\infty)} \frac{|\mathcal{A}(\boldsymbol{V},\boldsymbol{U})|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega_\infty)}}. \tag{4.47}$$

Proof. We will prove (4.46). Inequality (4.47) follows from (4.46) as in the proof of (4.34). For $\boldsymbol{U} \in \boldsymbol{H}^1_0(\Omega_\infty)$, we apply (4.35) to conclude

$$\|\boldsymbol{U}\|_{\boldsymbol{H}^1_0(\Omega_\infty)} \le C \sup_{\boldsymbol{V}\in\boldsymbol{H}^1_0(\Omega^c)} \frac{|\mathcal{A}(\boldsymbol{U},\boldsymbol{V}_0) + \mathcal{A}(\boldsymbol{U},\boldsymbol{V}_1)|}{\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega^c)}}$$

where we have decomposed $\boldsymbol{V} = \boldsymbol{V}_0 + \boldsymbol{V}_1$ with $\boldsymbol{V}_1$ satisfying $\boldsymbol{V}_1 = \boldsymbol{V}$ on Ω^c_∞, $\boldsymbol{V}_1 = \boldsymbol{0}$ on Ω_1 and

$$\mathcal{A}(\boldsymbol{\phi},\boldsymbol{V}_1) = 0, \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}^1_0(\Omega_\infty \setminus \overline{\Omega}_1). \tag{4.48}$$

The above problem is well posed. Indeed, for $\boldsymbol{\phi} \in \boldsymbol{H}^1_0(D)$, where D is any

subdomain of Ω_1^c

$$|\mathcal{A}(\boldsymbol{\phi},\boldsymbol{\phi})| = |d_0|^2|\mathcal{A}(\boldsymbol{\phi},\bar{d}_0^{-2}\boldsymbol{\phi})|$$
$$\geq -|d_0|^2 Im(\mathcal{A}(\boldsymbol{\phi},\bar{d}_0^{-2}\boldsymbol{\phi})) \geq \sigma_0 \min(1,k^2)\|\boldsymbol{\phi}\|^2_{\boldsymbol{H}^1(D)}.$$

It follows that from Lemma 1, that $\boldsymbol{V}_1$ is unique and satisfies

$$\|\boldsymbol{V}_1\|_{\boldsymbol{H}^1(\Omega_\infty\setminus\Omega_1)} \leq C\|\boldsymbol{V}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)} \leq C\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega_\infty\setminus\overline{\Omega}_1)}. \tag{4.49}$$

Let $\widetilde{\boldsymbol{U}}$ solve $\widetilde{\boldsymbol{U}} = \boldsymbol{U}$ on Ω_1 and

$$\mathcal{A}(\widetilde{\boldsymbol{U}},\boldsymbol{\phi}) = 0, \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega_1^c).$$

As above, from Lemma 1, $\widetilde{\boldsymbol{U}}$ is unique in $\boldsymbol{H}^1(\Omega_1^c)$ and satisfies

$$\|\widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^1(\Omega_1^c)} \leq C\|\boldsymbol{U}\|_{\boldsymbol{H}^{1/2}(\Gamma_1)} \leq C\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega_\infty)}. \tag{4.50}$$

We also define $\boldsymbol{U}_1$ by $\boldsymbol{U}_1 = \boldsymbol{U}$ on Γ_1, $\boldsymbol{U}_1 = \boldsymbol{0}$ on Ω_∞^c, and

$$\mathcal{A}(\boldsymbol{U}_1,\boldsymbol{\phi}) = 0, \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega_\infty \setminus \overline{\Omega}_1). \tag{4.51}$$

Again, from Lemma 1, $\boldsymbol{U}_1$ is unique in $\boldsymbol{H}^1(\Omega_\infty \setminus \overline{\Omega}_1)$ and we have that

$$\mathcal{A}(\boldsymbol{U}_1 - \widetilde{\boldsymbol{U}},\boldsymbol{\phi}) = 0, \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega_\infty \setminus \overline{\Omega}_1). \tag{4.52}$$

Since $\boldsymbol{V}_1 \in \boldsymbol{H}_0^1(\Omega_1^c)$ and $\boldsymbol{U} - \boldsymbol{U}_1 \in \boldsymbol{H}_0^1(\Omega_\infty \setminus \overline{\Omega}_1)$, (4.51) and (4.48) imply

$$|\mathcal{A}(\boldsymbol{U},\boldsymbol{V}_1)| = |\mathcal{A}(\boldsymbol{U}_1 - \widetilde{\boldsymbol{U}},\boldsymbol{V}_1)|$$
$$\leq C\|\boldsymbol{U}_1 - \widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^1(\Omega_1^c)}\|\boldsymbol{V}_1\|_{\boldsymbol{H}^1(\Omega_\infty\setminus\overline{\Omega}_1)}.$$

Because of (4.52) and Lemma 1

$$\|\boldsymbol{U}_1 - \widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^1(\Omega_\infty\setminus\overline{\Omega}_1)} \leq C\|\widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)}.$$

Also because of (4.50) and Lemma 1

$$\|\boldsymbol{U}_1 - \widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^1(\Omega_\infty^c)} = \|\widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^1(\Omega_\infty^c)} \leq C\|\widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)}.$$

Combining the last three inequalities, (4.49) and Lemma 3 gives

$$|\mathcal{A}(\boldsymbol{U}, \boldsymbol{V}_1)| \leq C\|\widetilde{\boldsymbol{U}}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)}\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega_\infty)} \leq Ce^{-\sigma_0 k_1 R_t}\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{V}\|_{\boldsymbol{H}^1(\Omega_\infty)}.$$

Thus,

$$\|\boldsymbol{U}\|_{H^1(\Omega_\infty)} \leq C \sup_{\boldsymbol{V}_0 \in H_0^1(\Omega_\infty)} \frac{|\mathcal{A}(\boldsymbol{U}, \boldsymbol{V}_0)|}{\|\boldsymbol{V}_0\|_{H^1(\Omega_\infty)}} + Ce^{-\sigma_0 k_1 R_t}\|\boldsymbol{U}\|_{\boldsymbol{H}^1(\Omega_\infty)}.$$

The inequality (4.46) follows taking R_t sufficiently large. This completes the proof of the theorem. □

We finally prove that the truncated elasticwave PML solution $\widetilde{\boldsymbol{u}}_t$ converges exponentially to the elasticwave solution in $\boldsymbol{H}^1(\Omega_0)$.

Theorem 11. *Assume that R_t is large enough that Theorem 10 holds. Let $\widetilde{\boldsymbol{u}}$ be in $\boldsymbol{H}^1(\Omega^c)$ and satisfy*

$$\mathcal{A}(\widetilde{\boldsymbol{u}}, \boldsymbol{\phi}) = 0, \quad \textit{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega^c).$$

Let $\widetilde{\boldsymbol{u}}_t$ be in $\boldsymbol{H}^1(\Omega_\infty)$ and satisfy

$$\mathcal{A}(\widetilde{\boldsymbol{u}}_t, \boldsymbol{\phi}) = 0, \quad \textit{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega_\infty)$$

and $\widetilde{\boldsymbol{u}}_t = \boldsymbol{u}$ on Γ and $\widetilde{\boldsymbol{u}}_t = \boldsymbol{0}$ on Γ_∞. Then

$$\|\widetilde{\boldsymbol{u}} - \widetilde{\boldsymbol{u}}_t\|_{\boldsymbol{H}^1(\Omega_\infty)} \leq Ce^{-\sigma_0 k_1 R_t}\|\boldsymbol{u}\|_{\boldsymbol{H}^{1/2}(\Gamma)} \tag{4.53}$$

and hence

$$\|\boldsymbol{u} - \widetilde{\boldsymbol{u}}_t\|_{\boldsymbol{H}^1(\Omega_0)} \leq Ce^{-\sigma_0 k_1 R_t}\|\boldsymbol{u}\|_{\boldsymbol{H}^{1/2}(\Gamma)}. \tag{4.54}$$

Proof. Since

$$\mathcal{A}(\widetilde{\boldsymbol{u}} - \widetilde{\boldsymbol{u}}_t, \boldsymbol{\phi}) = 0, \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}_0^1(\Omega_\infty),$$

Theorem 10, Lemma 1, Lemma 3 and Theorem 9 give

$$\begin{aligned}\|\widetilde{\boldsymbol{u}} - \widetilde{\boldsymbol{u}}_t\|_{\boldsymbol{H}^1(\Omega_\infty)} &\leq C\|\widetilde{\boldsymbol{u}}\|_{\boldsymbol{H}^{1/2}(\Gamma_\infty)} \leq Ce^{-\sigma_0 k_1 R_t}\|\widetilde{\boldsymbol{u}}\|_{\boldsymbol{H}^1(\Omega_\infty)} \\ &\leq Ce^{-\sigma_0 k_1 R_t}\|\boldsymbol{u}\|_{\boldsymbol{H}^{1/2}(\Gamma)}.\end{aligned}$$

Thus (4.53) follows and, since $\widetilde{\boldsymbol{u}} = \boldsymbol{u}$ on Ω_0, (4.54) also follows. This completes the proof of the theorem.

□

CHAPTER V

NUMERICAL APPROXIMATION OF THE ELASTIC WAVE PML PROBLEM

In this chapter, we analyze the Galerkin approximation for the elastic wave PML problem discussed in Chapter IV, and present the results of some numerical experiments.

As far as we know, the H^1-based *inf-sup* conditions of the previous chapter cannot, in general, be developed using a lower order perturbation argument based on Theorem 1 as was done for the Helmholtz equation in Chapter III. Accordingly, the classical finite element analysis for non-coercive problems [25] cannot be applied. In contrast, if we limit the size of the PML coefficient, then we shall see that it is possible to prove an *inf-sup* condition via perturbation and classical finite element analysis implies convergence of the finite element approximation.

Thus, the standard Galerkin approximation is well-posed (for small mesh sizes) provided that the PML damping parameter σ_0 is "small enough". In Section 5.3 we have briefly commented on two ways to avoid that restriction.

5.1. Analysis of the Galerkin approximation

For simplicity, we assume that $\partial\Omega$ is polyhedral and choose Γ_∞ so that Ω_∞ is also polyhedral. For a triangulation $\mathbb{T}_h$ of Ω_∞, let $\widetilde{\boldsymbol{V}}_h$ be a finite element space of continuous piecewise polynomial complex valued vector functions which vanish on Γ_∞. We further simplify by assuming that $\boldsymbol{g}$ coincides with a function in $\widetilde{\boldsymbol{V}}_h$ on Γ. The Galerkin approximation of $\widetilde{\boldsymbol{u}}_t$ is the function $\boldsymbol{u}_h \in \widetilde{\boldsymbol{V}}_h$ satisfying

$$\begin{aligned} \mathcal{B}(\boldsymbol{u}_h, \boldsymbol{\psi}_h) = 0 \quad &\text{for all } \boldsymbol{\psi}_h \in \boldsymbol{V}_h, \\ \boldsymbol{u}_h = \boldsymbol{g} \ &\text{on } \partial\Omega. \end{aligned} \tag{5.1}$$

Here $\boldsymbol{V}_h$ denotes the set of functions in $\widetilde{\boldsymbol{V}}_h$ which vanish on Γ and $\mathcal{B}(\boldsymbol{u},\boldsymbol{v})$ denotes the scaled form defined by

$$\mathcal{B}(\boldsymbol{u},\boldsymbol{v}) = \mathcal{A}(\boldsymbol{u}, \bar{d}^{-1}\boldsymbol{v}).$$

We note that Theorem 10 is obviously valid for the form $\mathcal{B}$ as well.

Our goal is to apply the so-called "Schatz finite element duality argument [25]" to show that, for sufficiently small mesh size h, the solution to (5.1) exists and is unique. Unfortunately to obtain this result, we shall have to put a smallness constraint on our PML function $\tilde{\sigma}$. To this end, we fix $\tilde{\sigma}_1$ to be a function in $C^2(\mathbb{R}^+)$ satisfying

$$\begin{aligned} &\widetilde{\sigma}_1(r) = 0 \quad \text{for } r \le r_0, \\ &\widetilde{\sigma}_1(r) = 1 \quad \text{for } r \ge r_1, \\ &\widetilde{\sigma}_1(r) \text{ increasing for } r \in (r_0, r_1). \end{aligned} \tag{5.2}$$

and set $\tilde{\sigma} = \sigma_0\tilde{\sigma}_1$. We start by proving a Gårding type inequality for the form $\mathcal{B}(\cdot,\cdot)$.

Lemma 4. *There exists $S_\gamma > 0$ and a positive constant C (depending on S_γ), such that whenever $\sigma_0 \le S_\gamma$,*

$$\|\boldsymbol{w}\|^2_{\boldsymbol{H}^1(\Omega_\infty)} \le C(|\mathcal{B}(\boldsymbol{w},\boldsymbol{w})| + \|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{w}\|_{\boldsymbol{L}^2(\Omega_\infty)}) \quad \textit{for all } \boldsymbol{w} \in \boldsymbol{H}^1(\Omega_\infty). \tag{5.3}$$

Proof. The sesquilinear form $\mathcal{A}$ can be rewritten as

$$\begin{aligned} \mathcal{A}(\boldsymbol{w},\boldsymbol{\psi}) = &\sum_{i=1}^{3} \int_{\Omega_\infty} \left(\left(\frac{\tilde{d}^2}{d}\boldsymbol{P} + d(\boldsymbol{I}-\boldsymbol{P}) \right) \boldsymbol{\nabla}\, \boldsymbol{w}_i \right) \cdot \boldsymbol{\nabla}\, \bar{\boldsymbol{\psi}}_i \, dx \\ &+ \gamma \int_{\Omega_\infty} (\tilde{d}^2 d)^{-1} \nabla \cdot \left(\left(\tilde{d}^2\boldsymbol{P} + \tilde{d}d(\boldsymbol{I}-\boldsymbol{P}) \right) \boldsymbol{w} \right) \nabla \cdot \left(\left(\tilde{d}^2\boldsymbol{P} + \tilde{d}d(\boldsymbol{I}-\boldsymbol{P}) \right) \bar{\boldsymbol{\psi}} \right) dx \\ &- k^2 \int_{\Omega_\infty} \tilde{d}^2 d \boldsymbol{w} \cdot \bar{\boldsymbol{\psi}} \, dx. \end{aligned}$$

Let $\boldsymbol{D}$ denote the matrix $(\tilde{d}/d - 1)\boldsymbol{P}$. Using the equalities

$$\left(\frac{\tilde{d}^2}{d}\boldsymbol{P} + d(\boldsymbol{I}-\boldsymbol{P}) \right) = d\left(\left(\frac{\tilde{d}}{d} + 1 \right) \boldsymbol{D} + \boldsymbol{I} \right) \text{ and } \left(\tilde{d}^2\boldsymbol{P} + \tilde{d}d(\boldsymbol{I}-\boldsymbol{P}) \right) = \tilde{d}d(\boldsymbol{D}+\boldsymbol{I})$$

gives

$$\begin{aligned}\mathcal{B}(\boldsymbol{w},\boldsymbol{\psi}) = \sum_{i=1}^{3}\int_{\Omega_\infty}\left(\left(\left(\frac{\tilde{d}}{d}+1\right)\boldsymbol{D}+\boldsymbol{I}\right)\boldsymbol{\nabla}\,\boldsymbol{w}_i\right)\cdot\boldsymbol{\nabla}\,\bar{\boldsymbol{\psi}}_i\,dx \\ +\gamma\int_{\Omega_\infty}\left(\nabla\cdot(\boldsymbol{D}+\boldsymbol{I})\boldsymbol{w}\right)\left(\nabla\cdot(\boldsymbol{D}+\boldsymbol{I})\bar{\boldsymbol{\psi}}\right)\,dx + L.O.\end{aligned} \tag{5.4}$$

Here we have used the notation "$L.O.$" to denote terms which have at least one undifferentiated component of $\boldsymbol{w}$ or $\boldsymbol{\psi}$ so that

$$|L.O.| \le C(\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(\Omega_\infty)} + \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{w}\|_{\boldsymbol{L}^2(\Omega_\infty)}). \tag{5.5}$$

A key point in our proof will be the fact that we can make various quantities involving $\boldsymbol{D}$ arbitrarily small by decreasing σ_0. We note that the coefficient $(\tilde{d}/d-1)$ is supported only in the transition region and that

$$\left|\frac{\tilde{d}}{d}-1\right| \le C\sigma_0. \tag{5.6}$$

Indeed,

$$\left|\frac{\tilde{d}}{d}-1\right| = \left|\frac{1+i\tilde{\sigma}}{1+i\sigma}-1\right| = \left|\frac{i(\tilde{\sigma}-\sigma)}{1+i\sigma}\right| \le \sigma_0 \max_r |\tilde{\sigma}_1(r)-\sigma_1(r)|\,. \tag{5.7}$$

We clearly have

$$\begin{aligned}\sum_{i=1}^{3}\int_{\Omega_\infty}\left(\left(\left(\frac{\tilde{d}}{d}+1\right)\boldsymbol{D}+\boldsymbol{I}\right)\boldsymbol{\nabla}\,\boldsymbol{w}_i\right)\cdot\boldsymbol{\nabla}\,\bar{\boldsymbol{\psi}}_i\,dx \\ = \sum_{i=1}^{3}\int_{\Omega_\infty}\boldsymbol{\nabla}\,\boldsymbol{w}_i\cdot\boldsymbol{\nabla}\,\bar{\boldsymbol{\psi}}_i\,dx + \sum_{i=1}^{3}\int_{\Omega_\infty}\left(\left(\left(\frac{\tilde{d}}{d}+1\right)\boldsymbol{D}\right)\boldsymbol{\nabla}\,\boldsymbol{w}_i\right)\cdot\boldsymbol{\nabla}\,\bar{\boldsymbol{\psi}}_i\,dx.\end{aligned} \tag{5.8}$$

In addition,

$$\nabla\cdot(\boldsymbol{D}\boldsymbol{w}) = tr(\boldsymbol{D}\boldsymbol{\nabla}\,\boldsymbol{w}) + div\boldsymbol{D}^t\cdot\boldsymbol{w},$$

where the divergence $div\boldsymbol{D}^t$ of the matrix function $\boldsymbol{D}^t$ is defined to be the vector

whose components are the divergences of the rows of $\boldsymbol{D}^t$. We can then rewrite

$$\begin{aligned}\gamma \int_{\Omega_\infty} (\nabla \cdot (\boldsymbol{D}+\boldsymbol{I})\boldsymbol{w}) \left(\nabla \cdot (\boldsymbol{D}+\boldsymbol{I})\bar{\boldsymbol{\psi}}\right)\, dx = \\ \gamma \int_{\Omega_\infty} \nabla \cdot \boldsymbol{w} \nabla \cdot \bar{\boldsymbol{\psi}}\, dx + \gamma \int_{\Omega_\infty} tr(\boldsymbol{D}\boldsymbol{\nabla}\, \boldsymbol{w})\ tr(\boldsymbol{D}\boldsymbol{\nabla}\, \bar{\boldsymbol{\psi}})\, dx \\ + \gamma \int_{\Omega_\infty} tr(\boldsymbol{D}\boldsymbol{\nabla}\, \boldsymbol{w})\ \nabla \cdot \bar{\boldsymbol{\psi}}\, dx + \gamma \int_{\Omega_\infty} \nabla \cdot \boldsymbol{w}\ tr(\boldsymbol{D}\boldsymbol{\nabla}\, \bar{\boldsymbol{\psi}})\, dx \\ + L.O.\end{aligned} \tag{5.9}$$

We now split the high-order terms from (5.8) and (5.9) in two parts, depending whether or not they involve the matrix $\boldsymbol{D}$. We introduce the forms

$$a_1(\boldsymbol{w}, \boldsymbol{\psi}) = \sum_{i=1}^{3} (\boldsymbol{\nabla}\, \boldsymbol{w}_i, \boldsymbol{\nabla}\, \boldsymbol{\psi}_i) + \gamma(\nabla \cdot \boldsymbol{w}, \nabla \cdot \boldsymbol{\psi}) + (\boldsymbol{w}, \boldsymbol{\psi}) \tag{5.10}$$

and

$$\begin{aligned} a_2(\boldsymbol{w}, \boldsymbol{\psi}) = \sum_{i=1}^{3} \int_{\Omega_\infty} \left(\left(\left(\frac{\tilde{d}}{d} + 1 \right) \boldsymbol{D} \right) \boldsymbol{\nabla}\, \boldsymbol{w}_i \right) \cdot \boldsymbol{\nabla}\, \bar{\boldsymbol{\psi}}_i\, dx \\ + \gamma \int_{\Omega_\infty} tr(\boldsymbol{D}\boldsymbol{\nabla}\, \boldsymbol{w})\ tr(\boldsymbol{D}\boldsymbol{\nabla}\, \bar{\boldsymbol{\psi}})\, dx \\ + \gamma \int_{\Omega_\infty} tr(\boldsymbol{D}\boldsymbol{\nabla}\, \boldsymbol{w})\ \nabla \cdot \bar{\boldsymbol{\psi}}\, dx + \gamma \int_{\Omega_\infty} \nabla \cdot \boldsymbol{w}\ tr(\boldsymbol{D}\boldsymbol{\nabla}\, \bar{\boldsymbol{\psi}})\, dx. \end{aligned} \tag{5.11}$$

Clearly, we have the equality

$$\mathcal{B}(\boldsymbol{u}, \boldsymbol{u}) = a_1(\boldsymbol{u}, \boldsymbol{u}) + a_2(\boldsymbol{u}, \boldsymbol{u}) + L.O.$$

Note that for any $\boldsymbol{u} \in \mathbf{H}^1(\Omega_\infty)$

$$\|\boldsymbol{u}\|^2_{\mathbf{H}^1(\Omega_\infty)} \leq |a_1(\boldsymbol{u}, \boldsymbol{u})| \text{ and } |a_2(\boldsymbol{u}, \boldsymbol{u})| \leq C_\gamma \sigma_0 \|\mathbf{u}\|^2_{\mathbf{H}^1(\Omega_1)}.$$

Making use of these bounds and inequality (5.5), we get

$$\begin{aligned}\|\boldsymbol{u}\|^2_{\mathbf{H}^1(\Omega_\infty)} &\leq |a_1(\boldsymbol{u},\boldsymbol{u})| \leq |\mathcal{B}(\boldsymbol{u},\boldsymbol{u})| + |a_2(\boldsymbol{u},\boldsymbol{u})| + |L.O.| \\ &\leq |\mathcal{B}(\boldsymbol{u},\boldsymbol{u})| + C_\gamma\sigma_0\|\boldsymbol{u}\|^2_{\boldsymbol{H}^1(\Omega_1)} + C_0\|\boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega_\infty)}.\end{aligned} \tag{5.12}$$

Taking $\sigma_0 \leq S_\gamma < \frac{1}{C_\gamma}$, for $\sigma_0 \leq S_\gamma$ we have

$$0 < (1 - C_\gamma S_\gamma) \leq (1 - C_\gamma\sigma_0)$$

and so

$$(1 - C_\gamma S_\gamma)\|\boldsymbol{u}\|^2_{\boldsymbol{H}^1(\Omega_\infty)} \leq (1 - C_\gamma\sigma_0)\|\boldsymbol{u}\|^2_{\boldsymbol{H}^1(\Omega_\infty)} \leq |\mathcal{B}(\boldsymbol{u},\boldsymbol{u})| + C_0\|\boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{u}\|_{\mathbf{L}^2(\Omega_\infty)}, \tag{5.13}$$

which proves inequality (5.3).

□

Remark 8. *It should be noted that, in addition to being independent of σ_0, the constants C_γ and C_0 (and consequently S_γ and C) in the proof above are also independent of the diameter of the region Ω_∞.*

In order to apply the duality argument, we shall also need the following regularity result.

Proposition 6. *There exists an $s > \frac{1}{2}$ and a constant $C_{reg} > 0$, such that for any $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega_\infty)$, the solution $\boldsymbol{w} \in \boldsymbol{H}^1_0(\Omega_\infty)$ of*

$$\mathcal{B}(\boldsymbol{\psi},\boldsymbol{w}) = (\boldsymbol{\psi},\boldsymbol{f}) \text{ for all } \boldsymbol{\psi} \in \boldsymbol{H}^1_0(\Omega_\infty),$$

is in $\boldsymbol{H}^{1+s}(\Omega_\infty)$ and satisfies

$$\|\boldsymbol{w}\|_{\boldsymbol{H}^{1+s}(\Omega_\infty)} \leq C_{reg}\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega_\infty)}.$$

Remark 9. *Full interior regularity of the solution $\boldsymbol{w}$ to this problem follows from the*

C^2 smoothness of the PML coefficients. Regularity near the boundary follows from known results for the regularity of solutions to the constant coefficient equation (4.1) *on polyhedral domains (see for example Theorems 2.3 (2D) and 4.5 (3D) of [21]).*

We are now ready to state the main result of this section, i.e., the unique solvability for the Galerkin approximation.

Theorem 12. *Assume that $\sigma_0 \leq S_\gamma$. There exists an $h_0 > 0$ such that whenever $h < h_0$, there is a unique solution $\boldsymbol{u}_h \in \widetilde{\boldsymbol{V}}_h$ of* (5.1) *and*

$$\|\widetilde{\boldsymbol{u}}_t - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} \leq C \inf \|\widetilde{\boldsymbol{u}}_t - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}. \tag{5.14}$$

The infimum above is taken over $\boldsymbol{v}_h \in \widetilde{\boldsymbol{V}}_h$ with $\boldsymbol{v}_h = \boldsymbol{g}$ on Γ.

Proof. Given Lemma 4 and Proposition 6, the theorem essentially follows from the well known argument given by Schatz ([25]) which we present here.

We start by applying the standard Nitsche trick ([22]) and showing that the error decays faster in the $\boldsymbol{L}^2$-norm than it does in the $\boldsymbol{H}^1$-norm. Suppose $\boldsymbol{u}_h$ satisfies (5.1) and consider $\boldsymbol{e}_h = \boldsymbol{u} - \boldsymbol{u}_h$. As a consequence of (4.47), there exists a unique solution $\boldsymbol{\theta}$ to

$$\mathcal{B}(\boldsymbol{v}, \boldsymbol{\theta}) = (\boldsymbol{v}, \mathbf{e}_h) \text{ for all } \boldsymbol{v} \in \mathbf{H}_0^1(\Omega_\infty).$$

Using the orthogonality property $\mathcal{B}(\boldsymbol{e}_h, \boldsymbol{v}_h) = 0$ for $\boldsymbol{v}_h \in \boldsymbol{V}_h$, we now have (for an appropriate choice of the approximation $\boldsymbol{\theta}_h \in \boldsymbol{V}_h$ of $\boldsymbol{\theta}$)):

$$\|\boldsymbol{e}_h\|^2_{\boldsymbol{L}^2(\Omega_\infty)} = \mathcal{B}(\boldsymbol{e}_h, \boldsymbol{\theta}) = \mathcal{B}(\boldsymbol{e}_h, \boldsymbol{\theta} - \boldsymbol{\theta}_h) \leq Ch^s \|\boldsymbol{\theta}\|_{\boldsymbol{H}^{1+s}(\Omega_\infty)} \|\boldsymbol{e}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}.$$

According to Proposition 6, $\|\boldsymbol{\theta}\|_{\mathbf{H}^{1+s}(\Omega_\infty)} \leq C_r eg \|\boldsymbol{e}_h\|_{\boldsymbol{L}^2(\Omega_\infty)}$, and consequently

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^2(\Omega_\infty)} \leq Ch^s \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}. \tag{5.15}$$

Existence and uniqueness of a solution $\mathbf{u}_h$ to (5.1) are now easy to see. Indeed, if

we suppose that $\mathbf{u}_h$ solves (5.1) with zero right-hand side, Lemma 5.3 and inequality (5.15) imply

$$\|\boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} \le C\|\boldsymbol{u}_h\|_{\mathbf{L}^2(\Omega_\infty)} \le Ch^s\|\boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)},$$

and consequently the only solution for small enough h is $\boldsymbol{u}_h \equiv 0$. Since the discrete problem (5.1) is a finite-dimensional square system, existence and uniqueness of its solution follow.

To show that the finite element error is quasi-optimal (i.e. (5.14) holds), consider an arbitrary $\boldsymbol{v}_h \in \widetilde{\boldsymbol{V}}_h$ with $\boldsymbol{v}_h = \boldsymbol{g}$ on the boundary Γ. Lemma 5.3 applied to $\boldsymbol{u}_h - \boldsymbol{v}_h$ asserts

$$\begin{aligned}\|\boldsymbol{u}_h - \boldsymbol{v}_h\|^2_{\boldsymbol{H}^1(\Omega_\infty)} &\le C(\mathcal{B}(\boldsymbol{u}_h - \boldsymbol{v}_h, \boldsymbol{u}_h - \boldsymbol{v}_h) + \|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{L}^2(\Omega_\infty)}) \\ &= C(\mathcal{B}(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u}_h - \boldsymbol{v}_h) + \|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}\|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{L}^2(\Omega_\infty)}),\end{aligned}$$

or

$$\|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} \le C(\|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} + \|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{L}^2(\Omega_\infty)}).$$

Triangle inequality and (5.15) now imply

$$\begin{aligned}\|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} &\le C(\|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} + \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{L}^2(\Omega_\infty)} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{L}^2(\Omega_\infty)}) \\ &\le C(\|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} + h^s\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}).\end{aligned}$$

Finally, we get

$$\begin{aligned}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} &\le \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} + \|\boldsymbol{u}_h - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} \\ &\le C(\|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega_\infty)} + h^s\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega_\infty)}),\end{aligned}$$

which (for sufficiently small h) gives the desired inequality (5.14).

□

5.2. Computational results

We illustrate the performance of the PML technique on a problem in the case of two spatial dimensions. We rig up a problem with known solution. Specifically, we take $\gamma = 3$ and $k = 1$ in equation (4.1) and consider the function

$$\boldsymbol{u}(x,y) = \boldsymbol{u}(r,\theta) = \boldsymbol{\nabla}\times\left(H_1^{(1)}(kr)\cos(\theta)\right) + \boldsymbol{\nabla}\left(H_1^{(1)}(k_1 r)\cos(\theta)\right). \tag{5.16}$$

It is clear that the above function satisfies (4.1)-(4.5). The first term is solenoidal while the second is irrotational and so both wave components are present.

We now consider approximating the solution of (4.1) with $\Omega = [-1,1]^2$ and $\boldsymbol{g} = \boldsymbol{u}$ (given by (5.16)) on Γ. By construction, the solution is just the function $\boldsymbol{u}$ given by (5.16).

To define the PML approximation explicitly, we take $r_0 = 3$, $r_1 = 4$ and $\sigma_0 = 1$. We truncate the domain so that $\Omega_\infty = (-5,5)^2 \setminus [-1,1]^2$. Although it is not clear that this choice of σ_0 satisfies the smallness assumption of our theorem, it nevertheless appears to work (as we shall see below). In all of our reported experiments, we shall compare the difference between the Galerkin solution $\boldsymbol{u}_h$ defined by (5.1) and $\boldsymbol{u}$ defined above.

Figure 5.1 gives a surface plot representation for the the real part of the second component of the exact solution and the finite element PML approximation. As suggested by the theory, the PML solution appears to be close to the exact solution in the inside of the PML layer and goes to zero quickly in the PML region. The effect of the transition region is further illustrated from the overhead view given in Figure 5.2, where we can clearly see the "cut-off" annulus $B_4 \setminus \overline{B}_3$.

To more precisely gauge the behavior of the method, we compute the norms of the error between $\boldsymbol{u}$ and $\boldsymbol{u}_h$ near $\partial\Omega$ (of course, this is the only meaningful computation

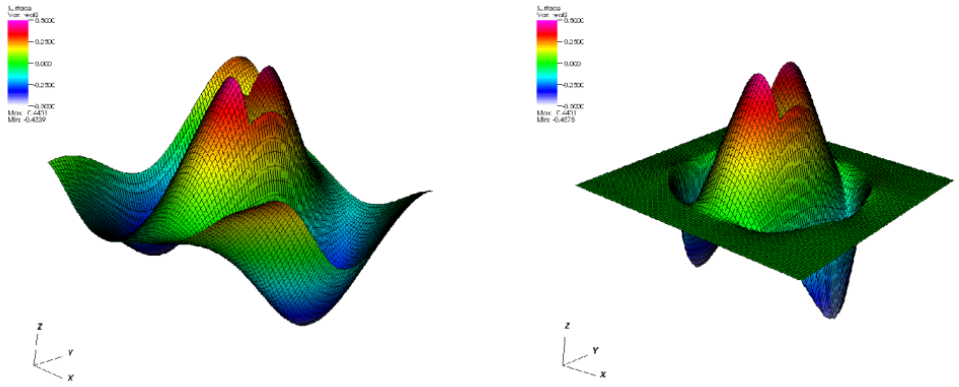

Fig. 5.1. The real part of the second component in the exact and the PML solutions.

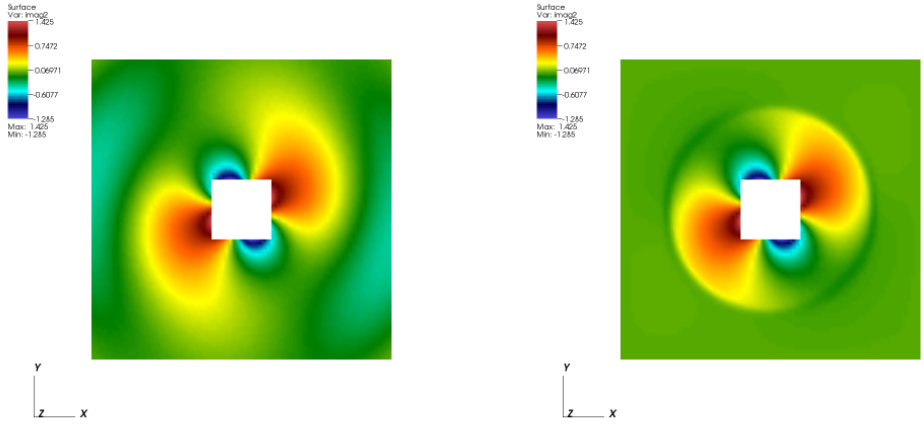

Fig. 5.2. The imaginary part of the second component in the exact and the PML solutions.

as $\boldsymbol{u}$ and $\boldsymbol{u}_h$ are significantly different in the PML region). Specifically, we report errors on $\Omega_* \equiv [-2,2]^2 \setminus [-1,1]^2$. Table 5.1 gives the $\boldsymbol{L}^2(\Omega_*)$ and $\boldsymbol{H}^1(\Omega_*)$ errors as a function of the mesh size. The $\boldsymbol{H}^1(\Omega_*)$ results clearly exhibit the expected first order of convergence. The $\boldsymbol{L}^2(\Omega_*)$ results show a convergence rate less than second order which is consistent with the fact that the domain Ω_∞ has re-entrant corners.

Table 5.1. L^2- and H^1- norms of the error

h	# dofs	$\boldsymbol{L}^2(\Omega_*)-$error	$\boldsymbol{H}^1(\Omega_*)-$error
$\frac{1}{2}$	1728	0.384539	1.990330
$\frac{1}{4}$	6528	0.274241	1.025252
$\frac{1}{8}$	25344	0.092068	0.482071
$\frac{1}{16}$	99840	0.029358	0.232628
$\frac{1}{32}$	396288	0.010079	0.114665

It is interesting to note that in the above calculations, we have not yet seen the pollution effect of the domain truncation come into play. This is not surprising as this error is exponentially small in the size of the domain and it appears that we have not yet made h small enough to see its effect.

5.3. Alternative approaches

The analysis given in the previous section imposes the smallness condition $\sigma_0 < S_\gamma$ to ensure well-posedness of the Galerkin approximation. It should be noted, that this condition might be artificial, as the numerical experiments suggest that the method also performs well for large values of σ_0 (i.e. $\sigma_0 = 10, 20, 50$). Nevertheless, in this section, we will briefly mention two alternative approaches that alleviate such a

restriction. The first is a negative-norm least-squares method based on the general framework for second order elliptic equations described in [5]. The second, which will be discussed in more detail in Chapter VI for the case of the Helmholtz equation, consists of coupling the PML for small σ_0 with a real scaling as the one outlined in Section 2.2.

We start by applying the least-squares technique for second order elliptic problems from [5]. That amounts to adding stabilizing terms to the continuous *inf-sup* condition, so that a discrete version of Theorem 10 holds (for the modified form).

Starting from inequality (4.46), integrating by parts and using the approximation properties of the finite element space, one arrives at the following inequality holding for functions $\boldsymbol{u} \in \boldsymbol{V}_h$

$$\begin{aligned} c\|\boldsymbol{u}\|^2_{\boldsymbol{H}^1_0(\Omega_\infty)} \leq \sup_{\boldsymbol{v}\in\boldsymbol{V}_h} \frac{|\mathcal{A}(\boldsymbol{u},\boldsymbol{v})|^2}{\|\boldsymbol{v}\|^2_{\boldsymbol{H}^1_0(\Omega_\infty)}} + \sum_{\tau\in\mathbb{T}_h} h_\tau^2 \| - \widetilde{d}^2 d(\widetilde{\boldsymbol{\Delta}}\boldsymbol{u} + \gamma\widetilde{\boldsymbol{\nabla}}\widetilde{\nabla}\cdot\boldsymbol{u} + k^2\boldsymbol{u})\|^2_{L^2(\tau)} \\ + \sum_{f\in\mathbb{F}_h} h_f \left(\sum_i \|[\![\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}\boldsymbol{\nabla}\, u_i\cdot\boldsymbol{n}]\!]\|^2_{L^2(f)} + \|[\![\gamma\widetilde{\nabla}\cdot\boldsymbol{u}]\!]\|^2_{L^2(f)} \right). \end{aligned} \tag{5.17}$$

In the above inequality, $\mathbb{F}_h$ denotes the set of interior faces of the triangulation $\mathbb{T}_h$ and h_τ (h_f) denotes the diameter of the particular element (face).

In view of the bound (5.17), a negative-norm least squares method can then be analyzed based on the form

$$\begin{aligned} \mathcal{B}(\boldsymbol{u},\boldsymbol{v}) = (\mathcal{L}_h\boldsymbol{u}, \mathcal{L}_h\boldsymbol{v})_{\boldsymbol{V}_h^*} \\ + \sum_{\tau\in\mathbb{T}_h} h_\tau^2 (\widetilde{d}^2 d(\widetilde{\boldsymbol{\Delta}}\boldsymbol{u} + \gamma\widetilde{\boldsymbol{\nabla}}\widetilde{\nabla}\cdot\boldsymbol{u} + k^2\boldsymbol{u}), \widetilde{d}^2 d(\widetilde{\boldsymbol{\Delta}}\boldsymbol{v} + \gamma\widetilde{\boldsymbol{\nabla}}\widetilde{\nabla}\cdot\boldsymbol{v} + k^2\boldsymbol{v}))_{L^2(\tau)} \\ + \sum_{f\in\mathbb{F}_h} h_f \left(\sum_i ([\![\widetilde{\boldsymbol{\nabla}} u_i\cdot\boldsymbol{n}]\!], [\![\widetilde{\boldsymbol{\nabla}} v_i\cdot\boldsymbol{n}]\!])_{L^2(f)} + ([\![\gamma\widetilde{\nabla}\cdot\boldsymbol{u}]\!], [\![\gamma\widetilde{\nabla}\cdot\boldsymbol{v}]\!])_{L^2(f)} \right), \end{aligned} \tag{5.18}$$

were the operator $\mathcal{L}_h : \boldsymbol{V}_h \mapsto \boldsymbol{V}_h^*$ is defined in terms of the form $\mathcal{A}(\cdot,\cdot)$, by

$$\langle \mathcal{L}_h \boldsymbol{u}, \boldsymbol{v} \rangle = \mathcal{A}(\boldsymbol{u}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in \boldsymbol{V}_h.$$

Namely, a discrete version of the variational problem (4.40) will be: find $\boldsymbol{u}_h \in \boldsymbol{V}_h$, such that for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$

$$\mathcal{B}(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\boldsymbol{F}, \mathcal{L}_h \boldsymbol{v}_h)_{\boldsymbol{V}_h^*} + \sum_{\tau \in \mathbb{T}_h} h_\tau^2 (\boldsymbol{F}, \widetilde{d}^2 d(\widetilde{\boldsymbol{\Delta}} \boldsymbol{v}_h + \gamma \widetilde{\boldsymbol{\nabla}} \widetilde{\nabla} \cdot \boldsymbol{v}_h + k^2 \boldsymbol{v}_h))_{L^2(\tau)}. \quad (5.19)$$

Note that the cell stabilizing terms on the right hand side in (5.19) make sense only for $\boldsymbol{F} \in \boldsymbol{L}^2(\Omega_\infty)$ and can be omitted if the data lack the necessary smoothness.

We shall not discuss the analysis of the method outlined above, as it falls into the framework of [5]. The implementation of the method, and specifically the $\boldsymbol{V}_h^*$-inner product, was already discussed in Section 3.4.2 of this dissertation.

Another approach is the coupling of a well-posed PML problem (i.e. with σ_0 taken sufficiently small) with a real scaling in the part of the domain past the PML transition region. Since the real scaling is just a change of variables, the new problem will still be well posed. The solution of the rescaled problem then exhibits faster decay than the solution to the original PML problem. This approach is considerably easier to implement than the negative-norm least-squares method described above, although it requires some tuning of the parameters (of the real scaling) for obtaining best results. Detailed results from numerical experiments applying this technique to the Helmholtz PML equation will be given in Section 6.3

CHAPTER VI

NON-SPHERICAL SCALING AND PML; EXPERIMENTS

So far in this text we have only considered spherical scaling and spherical PML. That is to say, the function $n(\boldsymbol{x})$, which essentially determined the scaling factor of a point $\boldsymbol{x} \in \mathbb{R}^d$, was always taken to be $n(\boldsymbol{x}) = \|\boldsymbol{x}\|_2$. Thus, the sets of points with the same scaling factor (i.e. the level sets of the function $\widetilde{\sigma}(n(\boldsymbol{x}))$) were either circles or spheres depending on the space dimension.

In this chapter, we will consider a scaling, whose level sets are the boundaries of squares/rectangles or cubes/cuboids. The reason for this is two-fold. First, a polyhedral underlying geometry is easier to deal with, for example when we need to align the mesh to the boundary of the transition region. To illustrate, consider the real scaling change of variables for the Laplace equation discussed in Section 2.2. As solutions to the (homogeneous) Laplace equation are smooth away from the boundary, the use of higher order finite elements might be desirable. This means that unless we have aligned the mesh to the discontinuities of (the k-th derivative of) the scaling factor $\widetilde{\sigma}(n(\boldsymbol{x}))$, we have to use a sufficiently smooth function $\widetilde{\sigma}(n)$ to ensure the optimal order of convergence. On the other hand, when the mesh is aligned with those discontinuities, we can use elements of arbitrary order using only continuous scaling factor, i.e. without having to modify the coefficients of the equation. The second reason for considering polyhedral geometries is to avoid the unnecessary restriction on the shape of the domain of interest, which the spherical scaling imposes. Indeed, consider a rectangular scatterer $\Omega = [-a, a] \times [-b, b] \subset \mathbb{R}^2$, whose sides a and b differ significantly in size ($a << b$), and suppose we are interested in the solution of our problem only at points whose distance from Ω is no bigger than $c > 0$ (i.e. "close to the scatterer"). It is desirable to be able to define a scaling that preserves the solution

only in the rectangle $(-(a+c),(a+c)) \times (-(b+c),(b+c))$ instead of the whole ball B_{b+c}, as this would significantly reduce the size of the computational domain and, consequently, the required degrees of freedom.

This chapter is divided into three sections. The first introduces a suitable choice of the function $n(\boldsymbol{x})$ and presents the results of some computational examples, when we apply that choice to the real scaling for the Laplace equation. In the second section, we briefly discuss the non-spherical PML. As we shall see, the justification of the method outlined there again imposes a smallness condition on σ_0. Although, just as in the case of the elastic wave equation from the previous chapter, this restriction seems to be artificial, it is desirable to have a way of overcoming it. One approach towards that goal is the coupling of the PML scaling with the real change of variables. This is suggested in the third and last section of this chapter.

6.1. Non-Spherical scaling

In this section, we will only be considering real scaling in two spatial dimensions. The function $\widetilde{\sigma}(n)$ will be obtained by shifting from the exponential scaling function given in (1.5). We shall consider the Laplace equation (1.1) and will follow the notation introduced in Chapter II.

We begin first with an example of a scaling that does not quite fall in the framework discussed so far. Namely, instead of scaling uniformly in both coordinates, as in (2.1), we examine the transformation

$$\widetilde{\boldsymbol{x}} = (\widetilde{d}(|x_1|)x_1, \widetilde{d}(|x_2|)x_2). \tag{6.1}$$

The reasoning behind this is clear, as such a transformation does not have the drawbacks of the spherical scaling discussed above. The discontinuities of the scaling

factors are located on straight lines, and the domain where no scaling takes place (i.e. $\widetilde{d}(|x_1|) = \widetilde{d}(|x_2|) = 1$) can easily be made rectangular, if we take different shifts in the definition of $\widetilde{\sigma}$ (and consequently $\widetilde{d}$) for the different coordinates.

The Jacobian of such transformation is given by

$$\boldsymbol{J} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix},$$

with $d_1 = 1 + \sigma(|x_1|)$ and $d_2 = 1 + \sigma(|x_2|)$. Formula (2.9) then gives the following expression for the scaled Laplace operator

$$\widetilde{\Delta} w = \frac{1}{d_1 d_2} \nabla \cdot \boldsymbol{\mu} \boldsymbol{\nabla} w,$$

with $\boldsymbol{\mu}$ given by

$$\boldsymbol{\mu} = \begin{pmatrix} \frac{d_2}{d_1} & 0 \\ 0 & \frac{d_1}{d_2} \end{pmatrix},$$

and thus we have, similarly to (2.18), the finite-domain variational problem: find $\widetilde{u}_t \in H_0^1(\Omega_\infty)$, such that

$$(\boldsymbol{\mu} \boldsymbol{\nabla} \widetilde{u}_t, \boldsymbol{\nabla} v) = (\det \boldsymbol{J} \widetilde{f}, v) \text{ for all } v \in H_0^1(\Omega_\infty). \tag{6.2}$$

While the above problem is certainly well posed, the anisotropy in the components of $\boldsymbol{\mu}$ poses difficulties in its numerical approximation. Indeed, unlike in the spherical case, the quantities d_1 and d_2 are no longer comparable in size. Consider, for example, the strip $S = (-r_0, r_0) \times \mathbb{R}$, where r_0 is the shift in the definition of $\widetilde{\sigma}$ (i.e. $\widetilde{\sigma}(r) = 0$ for $r < r_0$). For points inside this strip, no scaling is done on the first component, while the scaling on the second component (and consequently the value of d_2) can be arbitrarily large. This leads to an increasingly ill-conditioned coefficient at points away from the domain of interest, which in turn affects the condition num-

ber of the discrete system. As we shall see below, numerical experiments done with the above transformation, show the condition number of the discrete system (preconditioned with a standard multigrid V-cycle algorithm) growing like R^2 compared to $ln(R)^2$ for the spherical scaling. Here R denotes the diameter of the unscaled domain (i.e. the image of the computational domain under the chosen scaling).

The purpose of the above example was to justify the use of uniform scaling in all coordinates, i.e. a transformation as the one in (2.1). We can change the underlying geometry of such scaling with a proper choice of the function $n(\boldsymbol{x})$. To this end, instead of the ℓ_2-norm of a point $\boldsymbol{x} \in \mathbb{R}^2$, we shall use the ℓ_∞-norm as a base for our scaling

$$n(\boldsymbol{x}) = \|\boldsymbol{x}\|_\infty = \max_i\{|x_i|\}. \tag{6.3}$$

Note that with this choice, our transformation will only be piecewise smooth, independently of the smoothness of the chosen $\widetilde{\sigma}$, since all the expressions arising (e.g. the Jacobian) depend on which component of $\boldsymbol{x}$ has an absolute value equal to $\|\boldsymbol{x}\|_\infty$ and are discontinuous along the lines $|x_1| = |x_2|$. This means that one will always have to align the mesh with those lines of discontinuity. An example of a good triangulation of the computational domain for our standard test case is given in Figure 6.1. The red line segments show the boundary of the domain of interest. The discontinuities in the coefficient there can be corrected by taking sufficiently smooth sigma. The blue line segments show the discontinuities inherited from the choice of $n(\boldsymbol{x})$, which we can not correct.

The Jacobian $\boldsymbol{J}$ of the transformation is given as the following piecewise smooth

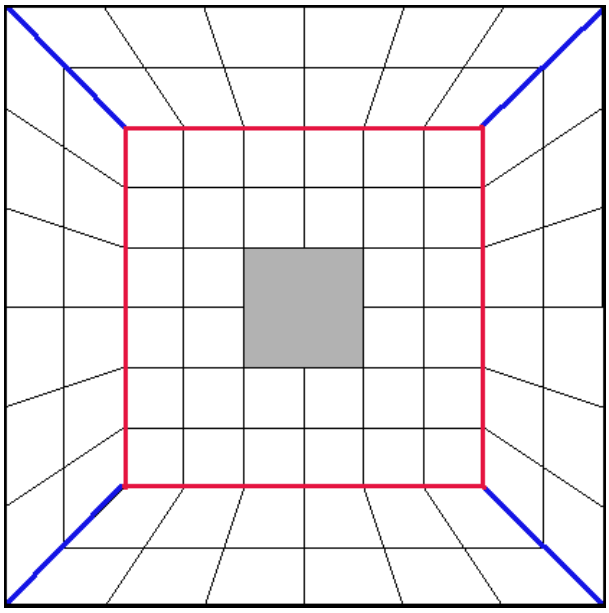

Fig. 6.1. A good grid for the infinity-norm scaling.

function

$$\boldsymbol{J} = \begin{cases} \begin{pmatrix} d & 0 \\ \widetilde{\sigma}'(|x_1|)x_2\, sgn(x_1) & \widetilde{d} \end{pmatrix} & \text{for} \quad |x_1| > |x_2|, \\ \begin{pmatrix} \widetilde{d} & \widetilde{\sigma}'(|x_2|)x_1\, sgn(x_2) \\ 0 & d \end{pmatrix} & \text{for} \quad |x_1| < |x_2|. \end{cases} \tag{6.4}$$

This gives the following formula for the matrix coefficient $\boldsymbol{\mu} = (\det \boldsymbol{J})\boldsymbol{J}^{-1}\boldsymbol{J}^{-T}$ in the expression for the scaled Laplacian (2.9).

$$\boldsymbol{\mu} = \begin{cases} \begin{pmatrix} \frac{\widetilde{d}}{d} & \frac{c_1}{d} \\ \frac{c_1}{d} & \frac{d}{\widetilde{d}} + \frac{c_1^2}{\widetilde{d}d} \end{pmatrix} & \text{for} \quad |x_1| > |x_2|, \\ \begin{pmatrix} \frac{d}{\widetilde{d}} + \frac{c_2^2}{\widetilde{d}d} & \frac{c_2}{d} \\ \frac{c_2}{d} & \frac{\widetilde{d}}{d} \end{pmatrix} & \text{for} \quad |x_1| < |x_2|. \end{cases} \tag{6.5}$$

Here $c_i = -\frac{\widetilde{\sigma}'(|x_i|)}{|x_i|}x_1x_2 = (\widetilde{d} - d)\frac{x_1x_2}{n(\boldsymbol{x})^2}$.

It can be seen that $\boldsymbol{\mu}$ satisfies inequality similar to (2.19) with absolute constants

and thus the infinity-norm scaling does not exhibit the degree of anisotropy that the componentwise scaling does. This is stated in the following proposition

Proposition 7. *The eigenvalues of* $\boldsymbol{\mu}$ *lie in the interval* $\left[\frac{1}{2}\frac{\widetilde{d}}{d}, 2\frac{d}{\widetilde{d}}\right]$.

Proof. Denote by a the ratio $\frac{d}{\widetilde{d}}$ and by s the quantity $\frac{x_1 x_2}{\|\boldsymbol{x}\|_\infty^2}$. We assume $|x_1| > |x_2|$ (as the other case is similar) and consider the matrix

$$a\boldsymbol{\mu} = \begin{pmatrix} 1 & (1-a)s \\ (1-a)s & a^2 + (1-a)^2 s^2 \end{pmatrix}.$$

Its eigenvalues $\lambda_1 \geq \lambda_2$ are given by the expression

$$\lambda_{1,2} = \frac{a^2 + (1-a)^2 s^2 + 1}{2} \pm \frac{1}{2}\sqrt{\left(a^2 + (1-a)^2 s^2 + 1\right)^2 - 4a^2}.$$

Clearly, for a fixed $a > 1$, λ_1 is an increasing function of $s^2 \in [0, 1]$. This gives the bounds

$$a^2 \leq \lambda_1 \leq a^2 + (a-1)^2 + 1 \leq 2a^2.$$

Analogously, one obtains for λ_2

$$\frac{1}{2} \leq \lambda_2 \leq 1.$$

Since the eigenvalues of $\boldsymbol{\mu}$ are $\frac{1}{a}\lambda_i$, the result follows. □

Table 6.1 shows the number of preconditioned conjugate gradient (PCG) iterations required for convergence and an estimate of the condition number of the preconditioned system for the two scalings discussed above. In both cases, the preconditioner used in the PCG algorithm is a standard V-cycle multigrid (c.f. Algorithm 3.1 in [9]) approximation of the inverse Laplacian. The results indicate that, as predicted, the uniform scaling based on the infinity-norm of a point results in a system considerably better conditioned than the one coming from componentwise scaling.

Table 6.1. Number of iterations and an estimated condition number for the infinity-norm scaling (infty) and the componentwise scaling (c/w) for three different scaling factors

		$R = 100$		$R = 200$		$R = 400$	
h	# dofs	infty	c/w	infty	c/w	infty	c/w
1	768	98	208	110	275	117	357
$\frac{1}{2}$	2976	127	623	153	1032	178	1603
$\frac{1}{4}$	11712	139	1000	166	1701	193	2790
$\frac{1}{8}$	46464	146	1276	175	2360	202	4020
$\frac{1}{16}$	185088	151	1498	181	2997	210	5317
$\frac{1}{32}$	738816	154	1629	184	3415	214	6523
cond. number		163	23274	236	121104	311	497693

Figure 6.2 presents a plot of the finite element approximation of the solution, $u(\boldsymbol{x}) = u(r, \theta) = \frac{\cos\theta}{r}$, to our usual test case, where the scatterer is $\Omega = [-1, 1]^2$, and the computational domain is $\Omega_\infty = (-5, 5)^2 \setminus \overline{\Omega}$. The scaling takes place in the part of the domain $\{\boldsymbol{x} \,|\, \|\boldsymbol{x}\|_\infty \geq 3\}$. One can clearly see that the approximation is close to the solution inside the domain of interest, as well as that the approximated solution is only piecewise smooth outside, as expected.

To illustrate the convergence, Table 6.2 shows the errors and the rate of convergence for the above test case when the approximation is done with bicubic finite elements. Note that with the appropriate mesh (i.e. the one shown on Figure 6.1), we obtain the optimal order of convergence, despite the fact that the scaling coefficient is only continuous.

Finally, we note that to allow for scatterers and, in general, domains of interest

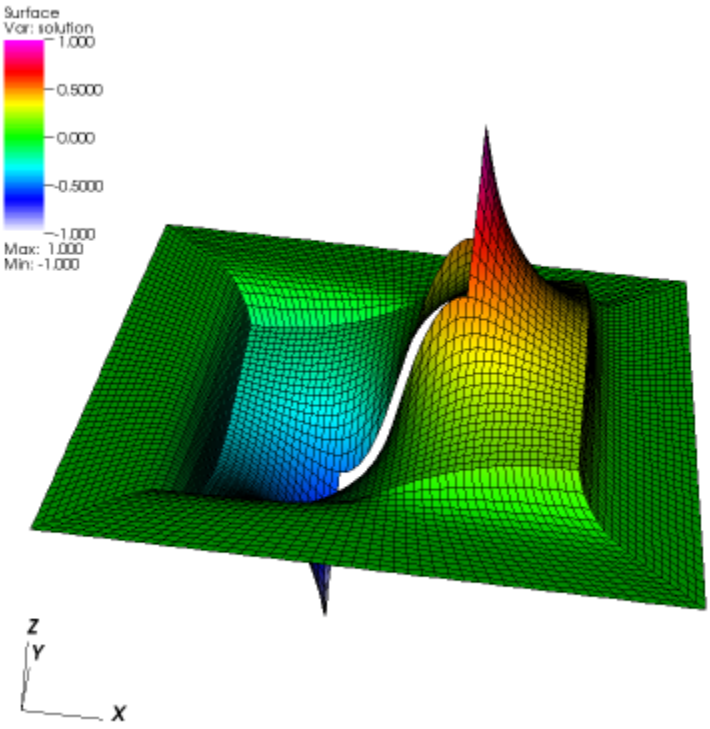

Fig. 6.2. Computed approximation (infinity-norm scaling).

Table 6.2. Errors in the domain of interest and the respective order of convergence

h	# dofs	L^2-error		H^1-error		L^∞-error	
1	768	4.89901e-03	-	3.67139e-02	-	6.62560e-03	-
$\frac{1}{2}$	2976	4.19660e-04	3.55	5.55492e-03	2.72	6.15797e-04	3.43
$\frac{1}{4}$	11712	2.64102e-05	3.99	7.44130e-04	2.90	5.89586e-05	3.38
$\frac{1}{8}$	46464	1.42007e-06	4.22	9.52298e-05	2.97	4.63213e-06	3.67
$\frac{1}{16}$	185088	1.34507e-07	3.40	1.19808e-05	2.99	3.19463e-07	3.86
$\frac{1}{32}$	738816	1.11253e-07	0.27	1.50203e-06	3.00	3.66809e-08	3.12

which are rectangular in shape (as opposed to square), one can consider uniform scaling based on the function

$$n(\boldsymbol{x}) = \max_i \{|x_i| + a_i\}, \tag{6.6}$$

where a_i, $i = 1 \dots d$ are non-negative numbers, at least one of which is zero.

To illustrate the performance of such scaling, we consider a test case, where the scatterer Ω is the rectangle $[-1.5, 1.5] \times [-0.5, 0.5] \subset \mathbb{R}^2$. We suppose the domain of interest consists of points which are at a distance of at most 1.5 units from the scatterer. This domain is contained in the rectangle $[-3, 3] \times [-2, 2] \subset \mathbb{R}^2$. We can describe the latter as $\{\boldsymbol{x} \,|\, n(\boldsymbol{x}) \leq 3\}$, where the function n is given as in (6.6) with $a_1 = 0$, $a_2 = 1$. Figure 6.3 presents a surface plot and an overhead view of the computed approximation (the exact solution $u(\boldsymbol{x})$ was taken as before). Note that the computation was done on the rectangular domain $\Omega_\infty = (-5, 5) \times (-4, 4) \setminus \Omega$ and we have reduced the required number of degrees of freedom by approximately 20% when compared to the infinity-norm scaling.

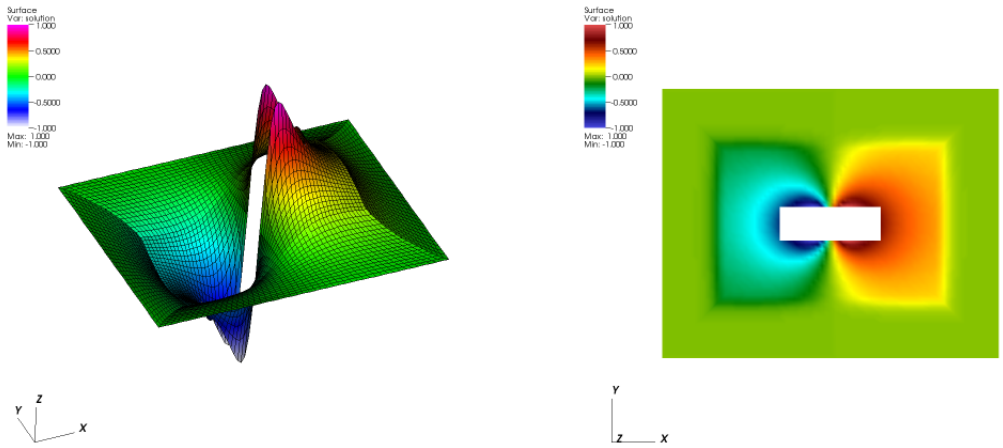

Fig. 6.3. Side (left) and overhead (right) view of the computed approximation.

6.2. Non-Spherical PML

In this section, we shall comment on the choice $n(\boldsymbol{x}) = \|\boldsymbol{x}\|_\infty$ for the Helmholtz PML equation. Recall that the PML scaling is obtained by taking $s = i$ in (2.1) and the formulas for the scaled operators. As usual when working with the PML scaling, we take the scaling factor $\widetilde{\sigma}(n)$ to be a shifted version of the function $\widetilde{\sigma}_{pml}$ given in (1.10).

Remark 10. *It should be noted that, with the above choice of $\widetilde{\sigma}$, a scaling of the type* (6.1) *is a reasonable choice. Indeed, the size of d_i is uniformly bounded in terms of σ_0, which is typically small. As such, the PML does not pose the difficulties in the numerical approximation discussed in the previous section. A complete analysis of such a PML was done by Pasciak and Kim in [19]. The analysis is more involved as one can not apply a perturbation argument similar to the one given in Chapter III, due to the fact that the resulting equation no longer has constant coefficients outside of a bounded domain.*

As the infinity-norm PML has no real advantages compared to the aforementioned

componentwise scaling, we shall only briefly discuss it here, setting the stage for the numerical experiments of the next section, where we give examples of the coupling of the real scaling and PML.

Following the development of the spherical PML in Chapter III, we start by defining the PML solution $\widetilde{u}(\boldsymbol{x})$ in terms of the series expansion for the original solution $u(\boldsymbol{x})$. In the two dimensional case, this means

$$\widetilde{u}(\boldsymbol{x}) = \begin{cases} u(\boldsymbol{x}) & \text{if } |\boldsymbol{x}| \leq r_0, \\ \displaystyle\sum_{n=1}^{\infty} H_n^{(1)}(k\widetilde{r}) p_n(\hat{\boldsymbol{x}}) & \text{if } |\boldsymbol{x}| \geq r_0. \end{cases} \tag{6.7}$$

This time we have $\widetilde{r} = \widetilde{d}(\|\boldsymbol{x}\|_\infty) r$. The functions p_n are defined on the unit circle and have the form $\alpha_n e^{in\theta} + \beta_n e^{-in\theta}$.

The function $\widetilde{u}(\boldsymbol{x})$ can be seen to satisfy the weak form of the PML equation

$$b(\widetilde{u}, v) = (f, v) \text{ for all } v \in H_0^1(\Omega^c), \tag{6.8}$$

where the form $b(\cdot,\cdot)$ is the same as (3.10), with the matrix coefficient $\boldsymbol{\mu}$ given by formula (6.5) from the previous section (with the above choice of the function $\widetilde{\sigma}(n)$ and the corresponding functions $\widetilde{d}(n) = 1 + i\widetilde{\sigma}(n)$ and $d(n) = \widetilde{d}(n) + n\widetilde{d}'(n)$).

It should be noted that the function $\widetilde{u}$ satisfies the strong form of the Helmholtz PML equation as well, since the discontinuities in the gradient $\boldsymbol{\nabla}\,\widetilde{u}$ and the coefficient $\boldsymbol{\mu}$ naturally cancel out. The justification of the equation itself is the same as for the spherical case (Theorem 3.9).

Note that, similarly to the spherical PML, the infinity-norm PML also results in a constant coefficient equation outside of a bounded domain containing the transition layer, which suggests the same approach to the analysis. Unfortunately, two problems arise. First, a uniqueness result, similar to Proposition 1, can not be shown using

the same techniques, and second, it is no longer true that $\Re(\boldsymbol{\mu})$ is uniformly positive definite for arbitrary values of σ_0. We can, however, state the following result.

Theorem 13. *Suppose that $w \equiv 0$ is the only solution of the variational problem*

$$b(w, v) = 0 \text{ for all } v \in H_0^1(\Omega^c). \tag{6.9}$$

Then there exists a constant $S > 0$ depending only on the size of the domain of interest and the transition region, such that for $\sigma_0 < S$, the variational problem (6.8) *is well-posed.*

Proof. Assumption (6.9) is essentially the result of Proposition 1. To proceed as in the analysis for the spherical case, we need a result similar to Proposition 2. Note that for $\sigma_0 = 0$ we have $\boldsymbol{\mu} \equiv \boldsymbol{I}$. Since the real part $\Re(\boldsymbol{\mu})$ and its eigenvalues depend continuously on the parameter σ_0, $\Re(\boldsymbol{\mu})$ will be a uniformly symmetric positive definite matrix for values of σ_0 in an interval $[0, S)$ for some $S > 0$. The proof of Proposition 2 then goes through without any modifications, as does the first part of Theorem 4.

It is clear that the constant S depends only on the choice of the function $\widetilde{\sigma}(n)$ and the range of values that $n(\boldsymbol{x})$ can take inside the PML transition region. □

Remark 11. *One might be able to prove assumption* (6.9) *using the techniques from [19], although in view of Remark 10, the interest in doing so will only be academic.*

Remark 12. *It should be noted that the value of the constant S in the theorem above is easily determined for a given particular case. It suffices to inspect the eigenvalues of $\Re(\boldsymbol{\mu})$ in the transition region $n(\boldsymbol{x}) \in (r_0, r_1)$.*

As in the case of the elastic wave equation, numerical experiments with the infinity-norm PML suggest that the smallness condition is artificial and the method

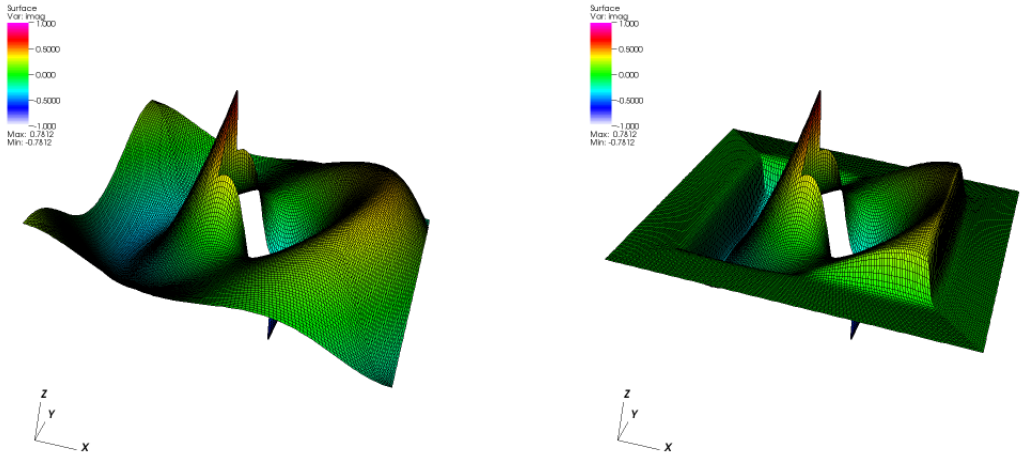

Fig. 6.4. The imaginary parts of the exact (left) and truncated (right) solutions to an exterior Helmholtz problem.

works in general. As noted in the the following section (cf. Figure 6.5), for the parameters in our usual test case, the eigenvalues of $\Re(\boldsymbol{\mu})$ are no longer positive for $\sigma_0 \geq 0.17$, yet the method performs well for larger values of σ_0. Figure 6.4 presents a plot of the imaginary part of the actual solution and its PML approximation, where σ_0 was taken equal to 1.

6.3. Combining PML with a real scaling

In this section, we will present some numerical experiments from coupling a PML scaling with a real change of variables outside the PML transition region. The motivation behind this approach is the fact that it is sometimes easier to theoretically justify the PML scaling, when the damping parameters σ_0 is small. We saw that in Chapter V, when discussing the elastic wave PML approximation, as well as in the previous section where we outlined the analysis of the infinity-norm PML applied to the Helmholtz equation. As any $\sigma_0 > 0$ introduces an exponential decay in the

solution to the PML problem, it is natural to attempt to speed-up the decay by compressing the space outside the transition region, where we are not interested in the solution anyway.

However, unlike the solutions to the Laplace equation, solutions to the Helmholtz and elastic wave equations are oscillating, and since the damping of the PML (for small σ_0) is not strong enough to quickly reduce the magnitude of the oscillations, this technique will not be as efficient. Compressing too many wave lengths in a unit size will require a sufficiently fine mesh to capture the oscillations in the solution, thus increasing the number of degrees of freedom needed to obtain a good approximation of the solution.

Nevertheless, as we shall discuss below, numerical experiments show that such an approach leads to a considerably better approximation than the PML with a small damping parameter alone.

In this section, once again, we shall work in the standard test case setting, where the scatterer is the square $\Omega = [-1,1]^2 \subset \mathbb{R}^2$ and the domain of interest is (as we are using the infinity-norm PML) inside the square $(-3,3)^2 \subset \mathbb{R}^2$. The boundary conditions on $\partial\Omega$ will be taken so that the exact solution is

$$u(\boldsymbol{x}) = u(r,\theta) = H_1^{(1)}(kr)\cos\theta,$$

for different values of the parameter k specified below.

For the parameters of the PML, we shall take $\widetilde{\sigma}(n) = \widetilde{\sigma}_{pml}(n-3)$, resulting in a transition region $T = \{\boldsymbol{x} \,|\, 3 \leq \|\boldsymbol{x}\|_\infty \leq 3+a\}$. Here the parameter a from the definition of $\widetilde{\sigma}_{pml}$ in (1.10) denotes the width of the transition region and will usually be taken $a = 1$. We shall experiment with different values of the parameter $\sigma_0 < 1$.

The computational domain will be the set $\Omega_\infty = (-R_\infty, R_\infty)^2 \setminus \Omega$. Unless otherwise specified, we shall take (as usual) $R_\infty = 5$.

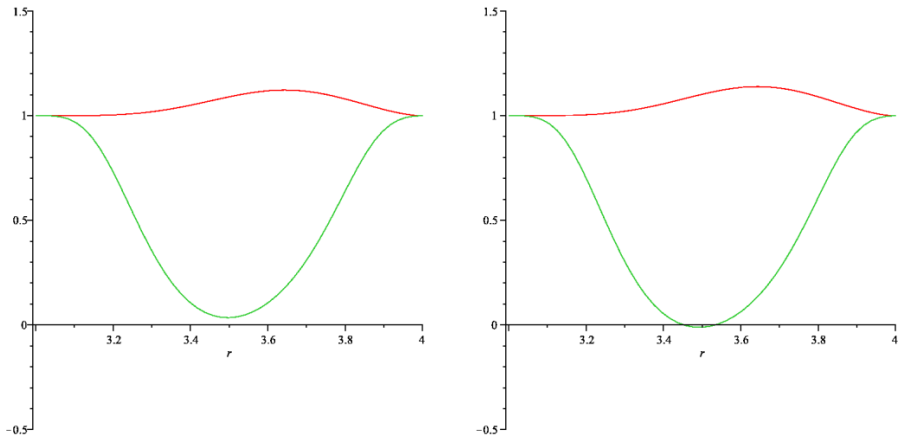

Fig. 6.5. Plot of the eigenvalues of $\Re(\boldsymbol{\mu})$ over the line $x_1 = x_2 = r \in (3, 4)$. Left - $\sigma_0 = 0.16$, Right - $\sigma_0 = 0.17$.

For the real scaling we take a function $\widetilde{\sigma}$, obtained in a similar manner from the function $\widetilde{\sigma}_{exp}$, which grows from 0 to $\frac{1}{\sigma_0} - 1$ over the interval $(3 + a, R_\infty)$. Thus the real scaling is done in the part of Ω_∞ past the PML transition region.

We first note that with the above parameters, the constant from the proof of Theorem 13 is $S \approx 0.166$. For a fixed point $\boldsymbol{x}$, the two eigenvalues of $\Re(\boldsymbol{\mu})$ deviate from 1, when increasing σ_0 and one of them can become negative when $\sigma_0 >= S$. A plot of the eigenvalues of $\Re(\boldsymbol{\mu})$ along the line segment $x_1 = x_2 = r \in (3, 4)$ for $\sigma_0 = 0.16$ and $\sigma_0 = 0.17$ is given on Figure 6.5.

To illustrate the benefit of introducing a real scaling to a well-posed PML problem, consider taking $\sigma_0 = 0.05$ in the above setting. The decay in the PML solution is so slow that the resulting approximation is visibly inaccurate. In fact, as shown in Table 6.3, the maximum pointwise error comes out to be approximately 50% (for the wavenumber $k = 1$). Introducing the real-scaling in the straight-forward manner above, decreases the error by three orders of magnitude.

Figure 6.6 shows the real parts of the exact solution, the PML approximation and the approximation resulting from coupling the PML with the real scaling in

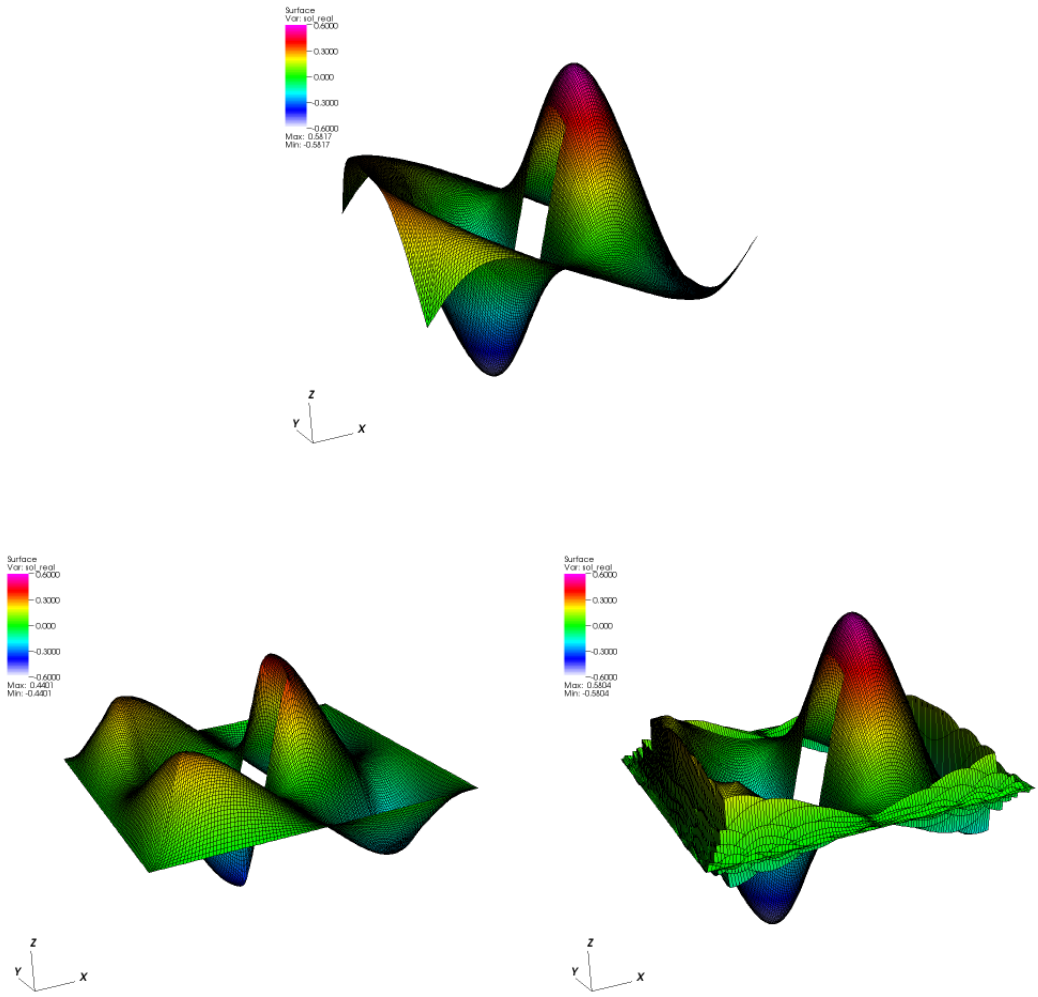

Fig. 6.6. Exact solution (top) and the results obtained with PML (left) and coupled scaling (right) for a small damping parameter.

Table 6.3. Norms of the error in the domain of interest for the two approaches; here $\sigma_0 = 0.05$

	L^2	H^1	L^∞
PML only	1.35550e+00	1.0794e+00	3.2583e-01
coupled scaling	5.90681e-04	4.91216e-04	1.37392e-04

the manner described above. Note that one can clearly see the oscillations in the solution to the coupled scaling equation. As we noted in the beginning of this section, when working with a limited number of degrees of freedom we can not expect to be able to always improve the approximation, when increasing the real scaling. In fact, if the real scaling parameter grows too fast over a small interval it is natural to expect the approximation error (within the limited degrees of freedom) to actually increase. Thus, there is some parameter tuning required to obtain the best possible approximation. Indeed, if the PML damping parameter σ_0 is taken too small, then the coefficients of the PML equation, and the solution itself, will be varying slowly inside the transition region. On the other hand, to compensate for the smallness of σ_0, one will need to increase the real scaling, resulting in more oscillations outside of that region. It is then natural to expect that one can get a better approximation when decreasing the size of the PML transition region (which we denoted by a above), thus allowing for a larger part of the computational domain to be used for the real scaling. Alternatively one might use a larger computational domain (i.e. increasing the value of R_∞), which will however increase the total number of degrees of freedom.

Table 6.4 presents the errors in the best possible approximations we were able to obtain when discretizing with 3.5 million degrees of freedom or less. We have applied the techniques described in this chapter to the Helmholtz PML equation with

Table 6.4. Best approximations for different values of the PML damping parameter

σ_0	real scaling	L^2-error	H^1-error	L^∞-error
5	no	4.75381e-09	1.08740e-06	1.40940e-08
1	no	1.58686e-08	1.09064e-06	3.02451e-08
0.1	yes ($a = 0.5$)	7.24748e-09	1.09253e-06	3.93642e-08
0.05	yes ($a = 0.5$, $R_\infty = 7$)	2.07709e-08	1.08873e-06	1.72234e-08
0.01	yes ($a = 0.25$, $R_\infty = 8$)	1.71997e-03	1.60201e-03	5.28631e-04

wavenumber $k = 1$, working with different values for the PML damping parameter σ_0 and applying additional real scaling to improve the performance when needed. The results serve to illustrate the following main points made in the discussion so far. First, the infinity-norm PML seems to work for arbitrary values of σ_0. Second, when working with a small σ_0, coupling the PML with a real scaling can lead to significant improvement of the resulting approximation. Third, some parameter tuning might be required to obtain the best possible approximation (compare the fourth line in Table 6.4 to the second line in Table 6.3). Lastly, although coupling with a real scaling substantially improves the approximation of the PML, for very small values of σ_0 the oscillations in the solution make the method less effective, compared to a PML with a larger damping parameter.

With regard to the last point above, we should note that for a fixed σ_0, the oscillations coming from solving an equation with an increased wave number, do not pose additional difficulty. The reason behind this is that the damping strength of the PML depends both on the size of σ_0 and k, and increasing the value of k increases the rate of decay of the PML solution.

Figure 6.7 presents plots of the PML approximation, the exact solution and the coupled scaling approximation the Helmholtz problem (in our usual test setting) with wavenumber $k = 3$. Here the PML damping parameter was taken to be $\sigma_0 = 0.1$. Note that while visually there is no detectable difference in the three functions inside the domain of interest, the addition of the real scaling to the PML decreased the error of the approximation by five orders of magnitude as shown in Table 6.5.

Table 6.5. Norms of the error in the domain of interest for the two approaches; here $\sigma_0 = 0.1$, $k = 3$

	L^2	H^1	L^∞
PML only	6.03772e-02	1.78356e-01	1.75885e-02
coupled scaling	5.56053e-07	7.19368e-06	1.64524e-07

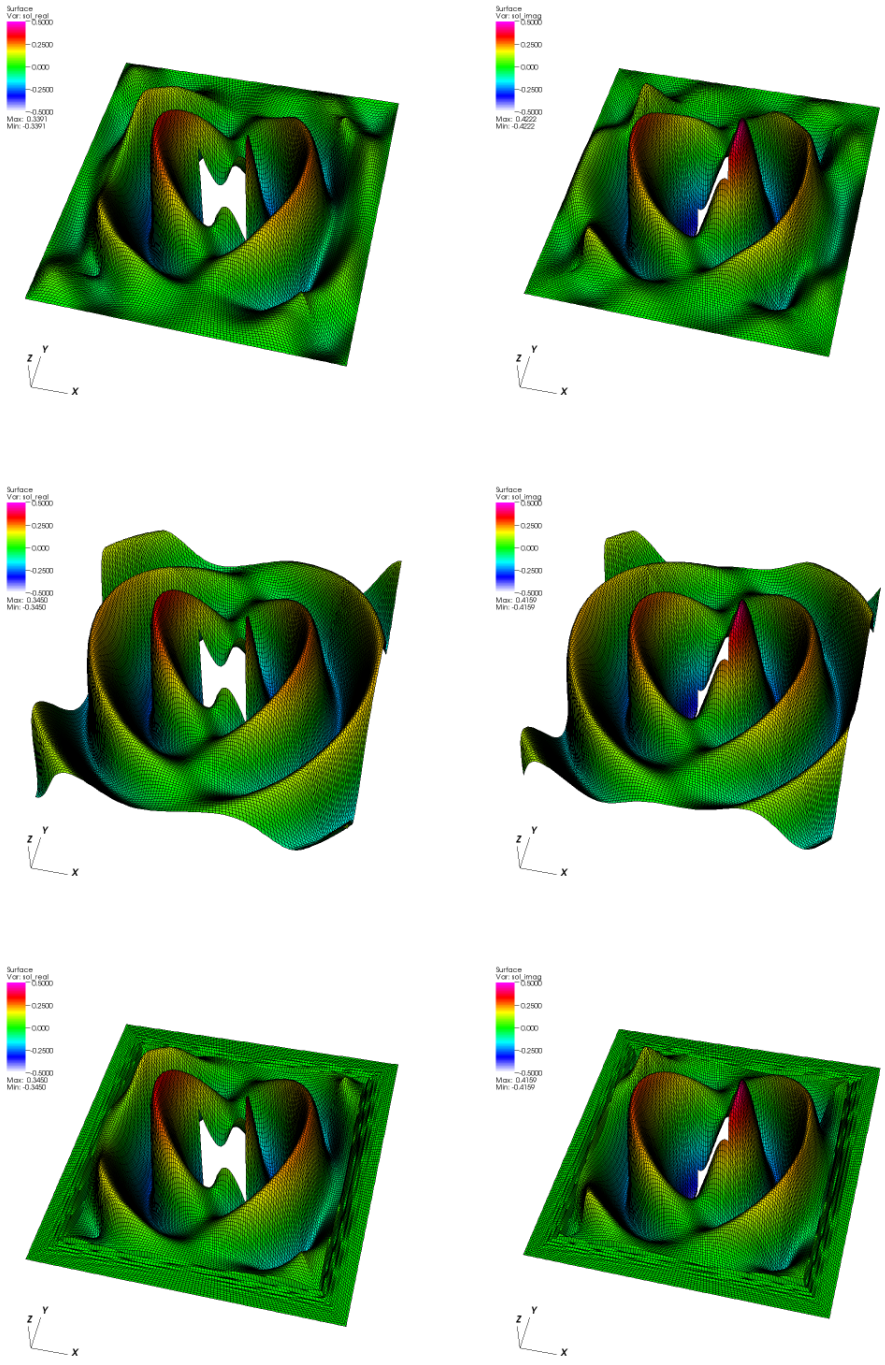

Fig. 6.7. The PML approximation (top), the exact solution (middle) and the coupled scaling approximation (bottom) for $k = 3$, $\sigma_0 = 0.1$. Real parts are shown on the left, imaginary parts - on the right.

CHAPTER VII

SUMMARY

In this chapter, we summarize the contents of this dissertation and briefly discuss possible future work.

7.1. Summary and conclusions

We have discussed some coordinate scaling techniques to modify certain boundary value problems posed on unbounded domains. The solutions of the transformed problems decay rapidly away from the region of interest while preserving the solution there. This allows an effective truncation of the new problems to relatively small domains, compared to the region of interest, thus introducing other new problems, suitable for numerical treatment using finite element approximation methods. Since the scaling technique essentially amounts to the introduction of variable coefficients into the original problem, it proves to be an easy method to implement for the numerical approximation of the solution.

For the exterior Laplace problem (1.1), we compared the performance of the real scaling technique with the existing methods of mesh-grading and adaptive mesh refinement for the solution of the problem. Tailored to the specific problem, the proposed method naturally performs better than a general adaptive refinement procedure in both accuracy (per degrees of freedom) and ease of implementation. Compared to the mesh grading technique, the real scaling approach is equally effective but has the advantage of being easier to implement and benefits from the underlying structured grid.

We introduced and examined the complex coordinate scaling for the acoustic scattering equation. As is well know, the complex stretched equation can be viewed

as a perfectly matched layer (PML) model for the truncation of the problem. We outlined the justification of the scaled equation and presented a proof of its well posedness. We discussed the implementation of the finite element discretization of the Helmholtz PML equation and the specific issues that arise. We also presented numerical results showing the efficiency of the PML method.

We then turned to the elastic wave scattering problem. We presented a detailed analysis of the elastic wave PML equation. While the results obtained were generalizations of the results for the acoustic wave PML, we note that the analysis is not a simple extension of the one presented for the simpler acoustic case. In particular, the *inf-sup* condition for the elastic wave case (e.g. Theorem 9) were not obtained using a lower order perturbation argument (e.g. in Theorem 3.23). This created additional difficulties in the analysis of the finite element discretization, where a restriction on the size of the PML damping parameter had to be imposed to derive a perturbed *inf-sup* condition (Lemma 4) and apply the standard finite element analysis. The restriction on the PML parameter seems to be artificial, since the method performs well in computational experiments regardless of its size.

Finally, we examined the use of different base functions for the scaling factor in both the real scaling and PML cases. We considered an infinity-norm based scaling allowing for a non-spherical (i.e. rectangular) underlying geometry. In the real case, we compared it to a Cartesian coordinate scaling (which, naturally, also allows for rectangular domains of interest) and discussed the advantages of the infinity-norm scaling. We outlined the analysis of the PML case, for which, once again, we had to impose a smallness restriction on the PML parameter. As a way to circumvent this restriction, we considered the coupling of real scaling with a PML with small damping parameter. We presented the results of extensive numerical experiments showing that this coupling leads to significant gains in the approximation quality than the PML

alone.

7.2. Future work

There are still many open questions that need need to be addressed concerning the techniques discussed. Extension of the results for time-dependent cases, as well as investigating efficient preconditioners for the solution of the resulting discrete problems, are among the natural directions for future analysis. Additionally, there are a few concrete problems that we plan to investigate in the future.

Consider, for example, the time-harmonic eddy-current problem. Its solutions, as shown in [1], decay like $O(\frac{1}{r^2})$ and their components (outside of the conductor) are harmonic functions. As such, the problem seems to be a perfect candidate for employing the real coordinate scaling technique.

We propose to couple the real space scaling approach with a suitable discretization to develop a fast and accurate method for the solving this problem. One possibility is to extend the negative-norm least-squares method for the time-harmonic Maxwell's equations from [4] to the eddy-current model. Another is to look at formulations where scalar potential of the electric or magnetic field is used outside of the conductor, making the problem (in this part of the domain) very similar to the exterior Laplace equation discussed Section 2.2. Both of those are reasonable discretization choices which we plan to investigate.

REFERENCES

[1] H. Ammari, A. Buffa, and J.-C. Nédélec, *A justification of eddy currents model for the Maxwell equations*, SIAM J. Appl. Math., 60 (2000) pp. 1805–1823.

[2] W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II—a general-purpose object-oriented finite element library*, ACM Trans. Math. Software, 33 (2007) Art. 24, 27.

[3] J.-P. Bérenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comput. Phys., 114 (1994) pp. 185–200.

[4] J. H. Bramble, T. V. Kolev, and J. E. Pasciak, *A least-squares approximation method for the time-harmonic Maxwell equations*, J. Numer. Math., 13 (2005) pp. 237–263.

[5] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *Least-squares for second order elliptic problems*, Comput. Meth. Appl. Mech. Engrg., 152 (1998) pp. 195–210.

[6] J. H. Bramble and J. E. Pasciak, *Analysis of a finite PML approximation for the three dimensional time-harmonic Maxwell and acoustic scattering problems*, Math. Comp., 76 (2007) pp. 597–614 (electronic).

[7] J. H. Bramble and J. E. Pasciak, *A note on the existence and uniqueness of solutions of frequency domain elastic wave problems: A priori estimates in H^1*, J. Math. Anal. Appl., 345 (2008) pp. 396–404.

[8] J. H. Bramble, J. E. Pasciak, and D. Trenev, *Analysis of a finite PML approximation for the three dimensional elastic wave scattering problem*, Math. Comp., submitted.

[9] J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, in Handbook of Numerical Analysis, P. G. Ciarlet and J. L. Lions, eds., Elsevier Science, Amsterdam, 2000, vol. VII, pp. 173–416.

[10] W. Chew and W. Weedon, *A 3D perfectly matched medium for modified Maxwell's equations with streched coordinates*, Microwave Opt. Techno. Lett., 13 (1994) pp. 599–604.

[11] F. Collino and P. Monk, *The perfectly matched layer in curvilinear coordinates*, SIAM J. Sci. Comp., 19 (1998) pp. 2061–2090.

[12] F. Collino and C. Tsogka, *Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media*, Geophysics, 66 (2001) pp. 294–307.

[13] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Applied Mathematical Sciences, 93. Springer-Verlag, Berlin, second edition, 1998.

[14] J. P. d. S. R. Gago, D. W. Kelly, O. C. Zienkiewicz, and I. Babuska, *A posteriori error analysis and adaptive processes in the finite element method. II. Adaptive mesh refinement*, Internat. J. Numer. Methods Engrg., 19 (1983) pp. 1621–1656.

[15] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*. Springer series in Computational Mathematics, 5. Springer-Verlag, Berlin, 1986.

[16] C. I. Goldstein, *The finite element method with nonuniform mesh sizes for unbounded domains*, Math. Comp., 36 (1981) pp. 387–404.

[17] F. D. Hastings, J. B. Schneider, and S. L. Broschat, *Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation*, The Journal of the Acoustical Society of America, 100 (1996) pp. 3061–3069, 1996.

[18] D. W. Kelly, J. P. d. S. R. Gago, O. C. Zienkiewicz, and I. Babuska *A posteriori error analysis and adaptive processes in the finite element method. I. Error analysis*, Internat. J. Numer. Methods Engrg., 19 (1983) pp. 1593–1619, 1983.

[19] S. Kim and J. E. Pasciak, *Analysis of a Cartesian PML approximation to acoustic scattering problems in* $\mathbb{R}^2$, submitted.

[20] J.-C. Nédélec, *Acoustic and Electromagnetic Equations*, Applied Mathematical Sciences, 144. Springer-Verlag, New York, 2001.

[21] S. Nicaise, *About the Lamé system in a polygonal or a polyhedral domain and a coupled problem between the Lamé system and the plate equation. I. Regularity of the solutions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., 19 (1992) pp. 327–361.

[22] J. Nitsche, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968) pp. 346–348.

[23] J. Peetre, *Espaces d'interpolation et théorèm de Soboleff*, Ann. Inst. Fourier, 16 (1966) pp. 279–317.

[24] M. Renardy and R. C. Rogers, *An Introduction to Partial Differential Equations*, Texts in Applied Mathematics, 13. Springer-Verlag, New York, second edition, 2004.

[25] A. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., 28 (1974) pp. 959–962.

[26] L. Tartar, *Topics in Nonlinear Analysis*, Math. d'Orsay, Univ. Paris-Sud, 1978.

[27] Y. Zheng and X. Huang, *Anisotropic perfectly matched layers for elastic waves in cartesian and curvilinear coordinates*, Research report, Massachusetts Institute of Technology, 2002, http://www-eaps.mit.edu/erl/research/report1/pdf2002/ZHENG.pdf

VITA

Dimitar Vasilev Trenev was born in Sofia, Bulgaria. He studied at Sofia University "St. Kliment Ohridski" from September 1998 and graduated in June 2002 with a Bachelor of Science degree in Mathematics. During the last two years of his studies, Dimitar interned as a computational software engineer at Sciant, Inc., currently a subsidiary of VMware, Inc. (NYSE: VMW). In the fall of 2002 he enrolled in the doctoral program in the Department of Mathematics, Texas A&M University. The current dissertation on scaling techniques for problems on unbounded domains was defended in April 2009. Dimitar can be contacted at:

Department of Mathematics

Texas A&M University

College Station, TX 77843-3368

U. S. A.

E-mail address: trenev@math.tamu.edu

The typist for this dissertation was Dimitar Trenev.